

Sichtfelderweiterung eines Laserscanners mittels einer Kamera für unbemannte Luftfahrzeuge

Von der Fakultät für Maschinenbau
der Technischen Universität Carolo-Wilhelmina zu Braunschweig
zur Erlangung der Würde
eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von: Stefan Krause
aus (Geburtsort): Brandenburg-Havel

eingereicht am: 11.02.2016
Mündliche Prüfung am: 19.05.2016

Gutachter: Prof. Dr.-Ing. Stefan Levedag
Institut für Flugsystemtechnik,
DLR Braunschweig

Prof. Dr.-Ing. Peter Hecker
Institut für Flugführung,
Technischen Universität Braunschweig

Danksagung

Diese Arbeit entstand während meiner Tätigkeit als Doktorand und wissenschaftlicher Mitarbeiter im Institut für Flugsystemtechnik in der Abteilung Unbemannte Luftfahrzeuge des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR) in Braunschweig.

Ich danke den Referenten Herrn Prof. Dr.-Ing. Stefan Levedag und Herrn Prof. Dr.-Ing. Peter Hecker für die Betreuung meiner Arbeit. Darüber hinaus danke ich dem Vorsitzenden meiner Prüfungskommission Herrn Prof. Dr.-Ing. Dirk Kügler.

Ebenso danke ich meinem Abteilungsleiter Jörg Steffen Dittrich für die Schaffung der erforderlichen zeitlichen und finanziellen Freiräume, welche die Durchführung der Dissertation erst ermöglichten.

Meinen Kollegen und Studenten in der Abteilung und im Institut danke ich für die Diskussionen und die Unterstützung. Besonderer Dank gilt meinem wissenschaftlichen Betreuer Dr.-Ing. Gordon Strickert für die Unterstützung, um sowohl inhaltlich als auch formal meine Arbeit zu einem erfolgreichen Abschluss zu bringen. Für zahlreichen Diskussionen und die Unterstützung bei den Flugversuchen danke ich Dr.-Ing. Sven Lorenz, Johann Dauer und Dr.-Ing. Franz Andert. Weiterer Dank gilt den Sicherheitspiloten Jörg Rößner und Michal Kislak-Schmidt, die durch ihr fliegerisches Können einen sicheren Flugbetrieb ermöglichten.

Für die Unterstützung bei der orthographisch Korrektur sowie der Beratung in der Endphase der Dissertation danke ich Dr.-Ing. Christoph Möhlenbrink, Rita Papenfuß und Nina Rose.

Ganz besondere Dankbarkeit geht an meine Eltern Waltraud und Hartmut Krause, die mich vom Beginn meiner schulischen Laufbahn, über das Studium, den Zeitraum der Promotion bis zum heute Tage begleitet, gestützt und fördert haben und mir mit Rat und Tat immer zur Seite standen.

Meine Freundin Anne Papenfuß möchte ich den tiefsten Dank aussprechen für Unterstützung und Beistand, welche sie jederzeit bereit war zu geben. Darüber hinaus bedanke ich mich für das Aushalten meiner nicht immer netten Launen, auch in Urlaub und Freizeit.

Allen meinen Freunden und meine ganzen Familie danke ich für die Ausdauer, Ruhe, Geduld sowie Aufmunterungen, die sie mir über die Jahre entgegengebracht haben.

Stefan Krause

Kurzfassung

Nutzungskonzepte für kleine unbemannte Luftfahrzeuge beschreiben oftmals eine Verwendung in Bodennähe, nah an Objekten und innerhalb einer unbekanntem Hinderniskulisse. Randbedingungen wie ein Betrieb außerhalb der Sicht des Operators und die Fortführung eines Einsatzes auch bei gestörter Kommunikation erfordern, dass die Luftfahrzeuge auch ohne direkte Anweisungen selbständig und sicher operieren können. Einsätze in einem unbekanntem Gelände setzen zusätzlich die automatische Detektion von Hindernissen voraus, um Kollisionen zu verhindern. Für kleine unbemannte Luftfahrzeuge mit begrenzter Nutzlast stehen für die automatische Hindernisdetektion nur wenige Sensoren mit unterschiedlichen Nachteilen, wie begrenztem Auflösungsvermögen, Sichtfeld oder geringer Signifikanz der Daten, zur Verfügung. Aktuelle Sensoren können alle Anforderungen, welche vom Luftfahrzeug ausgehen oder an eine Umweltwahrnehmung gestellt werden, allein oft nicht erfüllen.

Diese Arbeit präsentiert eine automatische Detektion von Hindernissen auf Basis einer Fusion monokularer Kamera- und monogoner Laserscannerdaten, welche eine Umgebungswahrnehmung bereitstellt, die gegenüber dem alleinigen Einsatz des Laserscanners erweitert ist und eine umfassendere Sicht auf die Umwelt ermöglicht. Darüber hinaus wird eine automatische Kalibrierung der beiden Sensoren zueinander sowie Verfahren für die Integration von kontinuierlichen Entfernungsmessungen in ein diskretes Weltmodell vorgestellt. Das Ergebnis dieses Vorgehen ist ein metrischen Weltmodell, welches den Raum, der das Luftfahrzeug umgibt, diskretisiert und pro Zelle beschreibt, ob der jeweilige Raum durch ein Hindernis belegt ist und somit nicht gequert werden kann.

Voraussetzung für die Fusion ist eine Bestimmung der extrinsischen Parameter zwischen der monokularen Kamera und dem monogonen Laserscanner. In der Arbeit wird ein automatisches Verfahren vorgestellt, das neben den Entfernungs- auch die Intensitätsmessungen des Laserscanners verwendet, um Fehlmessungen des Laserscanners welche auf Kantentreffern basieren, zu detektieren.

Als weitere Voraussetzung werden Lösungsansätze für die Integration von Entfernungsmessungen in ein diskretes Weltmodell vorgestellt, welche eine ungewollte Freizeichnung von Zellen verhindert, die als Folge einer Diskretisierung nur teilweise belegt sind. Diese Verfahren werden bezüglich ihrer Rechenperformanz verglichen und ein Ansatz mit der besten Eignung für den Betrieb auf einem kleinen unbemannten Luftfahrzeug ausgewählt.

Die dargelegte Fusion ist zweistufig aufgebaut und nutzt eine kooperative

Fusion, auf die eine konkurrierende Fusion folgt. Die kooperative Fusion der Sensordaten generiert dreidimensionale vereinfachte planare Approximationen der betrachteten Hindernisoberflächen, wobei deren Lage und Position auf Lasermessungen und die Abmessungen auf monokularen Kameradaten beruhen. Die Annahme einer planaren Oberflächenapproximation bildet die tatsächlichen Hindernisse nur sehr grob ab. Aus diesem Grund werden mehrere Approximationen, die zu unterschiedlichen Zeitpunkten und Blickwinkeln von den Hindernissen generiert wurden, in einem metrischen Weltmodell, einem OcTree, unter Verwendung eines probabilistischen Sensormodells konkurrierend fusioniert, um eine realistischere Repräsentation der Hindernisse zu erreichen.

Der Fokus bei der Entwicklung lag auf kleinen unbemannten Luftfahrzeugen und den für diese Klasse von Vehikeln zur Verfügung stehenden Sensoren. Die umgesetzten Algorithmen sind für den Betrieb an Bord des Vehikels vorgesehen und im Hinblick auf die begrenzte Rechenleistung optimiert.

Validiert wurden die Verfahren in Simulationen und mittels realer Testdaten, die mit einem unbemannten Helikopter als Sensorträger erstellt wurden. Die Ergebnisse zeigen eine erfolgreiche Erweiterung der Hinderniswahrnehmung, welche für das dargelegte Beispiel um 140% über die Sicht des Lasers hinausgeht. Der Ansatz ermöglicht damit einen Betrieb des unbemannten Vehikels in einer unbekanntem Hinderniskulisse, der sicherer ist als bei alleiniger Nutzung eines Einzelsensors.

Abstract

Utilization concepts for small unmanned aerial vehicles usually describe operations at near-ground level, close to objects and in the vicinity of unknown obstacles. Boundary conditions such as operations outside the operator's line of sight, as well as potential loss of communication require that the aircraft be capable of operating independently and safely without direct instructions. In addition, operations in an unknown terrain also require the ability to detect obstacles reliably, to prevent collisions. For small unmanned aerial vehicles with a limited payload, only few sensors are available to enable automatic obstacle detection, and those have various disadvantages, such as limited resolution, field of view or their data have an insufficient significance. Therefore, current sensors can often not fulfil all requirements, which emanate from the aerial vehicle, or are imposed by environment awareness.

This work presents an automatic three dimensional detection of obstacles based on a fusion of monocular camera and monogonal laser scanner data that provides an improved environmental awareness compared to the use of a laser scanner alone. In addition, methods for an automatic calibration of the two sensors to each other as well as for the integration of continuous distance measurements in a discrete world model are presented. The result of this collective approach is a metric world model, which could be used for automatic navigation of the small unmanned aerial vehicle in an unknown environment. In this model, the surrounding space of the aircraft discretizes where each, cell contains information whether it is occupied by an obstacle and thus cannot be crossed.

Requirement for application of the sensor data fusion is the knowledge of the extrinsic parameter between both sensors. Therefore, an approach for the automatic determination of these parameters between is introduced, which uses the intensity measurements of the laser scanner in addition to the distance measurements. The use of the intensity measurements enables an automatic detection of the calibration object and detects aberrations at the laser data, which base on edge hits.

A further component for the sensor data fusion and thus environment awareness is a correct integration of distance measurements in the metric world model. For the integration the continuous distance measurements need to be discretized. This leads to the problem that based on the resolution of the discretization some partly occupied cells are erroneously wrongly shown free. To solve this and to prevent collisions, three different solutions are presented which prevent an unintentional neglect of occupied cells in the world model.

The presented sensor data fusion approach is structured in two-levels and uses a collaborative fusion, followed by a competing fusion. The collaborative fusion of sensor data generates three-dimensional, simplified planar approximations of obstacle surfaces. The location and position of the approximation are based on laser measurements and the dimensions are based on monocular camera data. However, assuming a planar surface approximation limits the reproduction of the real shape of natural obstacles. For this reason, several approximations generated over different times and angles of view on the obstacles are fused in competitive manner within the metric world model. For this a probabilistic sensor model has been used to achieve a more realistic representation of the obstacles.

The approach developed in this work focuses on small unmanned aerial vehicles and limited capabilities of the sensors available for this class of vehicle. The algorithms used are intended for operating onboard of the vehicle and have been optimized to take into account of its limited computing power.

The procedures and solutions were validated in simulations and with real-world test data, created by using an unmanned helicopter as a sensor carrier. The results show a successful enhancement of the obstacle perception which in the presented example exceeds the laser scanner view about 140%. Therefore, it is safer to operate an unmanned vehicle in an unknown obstacle environment with the presented sensor data fusion approach of this work, as when employing a single sensor.

Inhaltsverzeichnis

Tabellenverzeichnis	iii
Abbildungsverzeichnis	iv
Abkürzungsverzeichnis	viii
Symbolverzeichnis	ix
1 Einleitung	1
1.1 Motivation	1
1.2 Automatische Hindernisdetektion	3
1.2.1 Beschaffung von Umgebungsinformationen	4
1.2.2 Kooperative Fusion	6
1.2.3 Konkurrierende Fusion	7
1.3 Aufbau der Arbeit	9
2 Grundlagen	11
2.1 Sensordatenfusion	11
2.2 Geometrische Konventionen und Grundlagen	15
2.2.1 Verwendete Koordinatensysteme	15
2.2.2 Koordinatentransformation	16
2.2.3 Verwendete Koordinatentransformation	18
2.2.4 Pose von Sensorträger und Kamera	20
2.3 Sensoren zur Umgebungserkennung	20
2.3.1 Überblick	21
2.3.2 LiDAR	22
2.3.3 Monokulare Kamera	37
2.4 Segmentierung	41
2.4.1 Überblick	41
2.4.2 Kantendetektion	43
2.5 Hindernisrepräsentation	48
2.5.1 Überblick	48
2.5.2 Belegtheitsgitter	50
2.5.3 Datenstruktur	54

3	Voraussetzungen für multisensorielle Hindernisdetektion	60
3.1	Extrinsische Kalibrierung	60
3.1.1	Überblick	62
3.1.2	Ermittlung der extrinsischen Parameter	63
3.1.3	Kantentreffer als Fehlerquelle	65
3.1.4	Erweitertes Kalibrierobjekt	69
3.1.5	Validierung	73
3.2	Diskretisierung der Umwelt	81
3.2.1	Problemdarstellung	83
3.2.2	Reduktion ungewollter Freizeichnung mittels großer Punktmengen	86
3.2.3	Reduktion ungewollter Freizeichnung mittels OcTree	87
3.2.4	Bedingte Freigabe belegter Zelle	88
3.2.5	Validierung	92
4	Erweiterung einer monogonen Lasersicht mittels Kamera	103
4.1	Stand der Technik	103
4.1.1	Homogene Konfigurationen	104
4.1.2	Heterogene Konfigurationen	105
4.2	Konzept	107
4.3	Vorverarbeitung	117
4.3.1	Vorverarbeitung des Kamerabildes	117
4.3.2	Vorverarbeitung der Laserdaten	120
4.4	Kooperative Fusion	122
4.4.1	Segmentierung	123
4.4.2	Anpassung der LiDAR-Auflösung	131
4.4.3	Oberflächenapproximation	135
4.5	Konkurrierende Fusion	139
4.5.1	Interpretation von Sensordaten	141
4.5.2	Rekonstruktion einer Objektoberfläche	148
4.6	Validierung der Sichtfelderweiterung	153
4.6.1	Versuchshardware	153
4.6.2	Umgebungskartierung mit Sichtfelderweiterung . . .	157
5	Zusammenfassung und Ausblick	174

Tabellenverzeichnis

2.1	Beispiele für die Remissionsgrade verschiedener Materialien und Farben [82, S. 8].	25
3.1	Parameter zur Unterdrückung der ungewollten Freizeichnung.	94
4.1	Kennzahlen der midiARTIS-Konfiguration.	155
4.2	Kennzahlen der Sensorkonfigurationen.	157
4.3	Kontingenztafel der verwendeten Signalentdeckungstheorie.	167

Abbildungsverzeichnis

1.1	Potenzielle Nutzungsszenarien für UA.	2
1.2	Blockdiagramm der Sichtfelderweiterung.	5
1.3	Approximation einer Objektoberflächenebene mittels unabhängiger Lasermessungen und Kamerabildern.	7
2.1	lokales Welt-, körperfestes Träger- und Sensorkoordinatensystem	16
2.2	Transformationen zwischen Welt-, Träger-, Kamera- und Laserkoordinatensystem	19
2.3	Strahlendivergenz bei Laserstrahlen	23
2.4	Prinzipieller Aufbau für eine Laufzeitenmessung	24
2.5	Ablenkung eines Laserstrahls mit zwei rotatorisch gelagerten Spiegeln.	27
2.6	Bewegungsbereinigte Laserscannermessungen durch eine Transformation der Daten in das lokale Weltkoordinatensystem.	29
2.7	Edge Effect: ein ausgesendeter Laserpuls generiert durch die simultane Beleuchtung von zwei Objekten, zwei Echos.	30
2.8	Kombiniertes Sensormodell nach Fox et al. [28].	34
2.9	Inverses Sensormodell abgeleitet von einer idealen Sensormessung, nach Elfes [24].	35
2.10	Inverses Sensormodell unter Berücksichtigung der von Fox et al. [28] dargelegten Fehlerquellen.	36
2.11	Lochkameramodell	37
2.12	Exemplarische Anwendung eines Sobel-Operators.	46
2.13	Exemplarisches Histogramm mit konvexer Hülle zur Herleitung der Canny-Operator Hystereseparameter.	48
2.14	Zweidimensionales Belegtheitsgitter, in das inkrementell Messungen \mathbf{p} eingefügt werden. Nach Scholz [81].	54
2.15	Geordneter, gewurzelter, Binärbaum der Höhe 3.	55
2.16	Rekursive Dekomposition eines dreidimensionalen Raumes mittels eines OcTree der Höhe drei.	57
2.17	Rekursive Dekomposition eines zweidimensionalen Objekts mittels eines QuadTree ¹ der Höhe vier.	58

3.1	Aufbau einer extrinsischen Kalibrierung zwischen monokularer Kamera c und monogonem LiDAR l	64
3.2	Vermessung einer Objektausdehnung $d(\mathbf{D}_0, \mathbf{D}_1)$	65
3.3	LiDAR Entfernungs- und Intensitätsmessungen eines bekannten Objekts [54].	66
3.4	Fehler bei der Vermessung eines bekannten Objekts abhängig von dessen Oberflächenreflektivität, nach [54].	67
3.5	Kantentreffer bei „First-Echo“-Auswertung	67
3.6	Fehler bei der Vermessung eines bekannten Objektes abhängig von dessen Oberflächenreflektivität, nach [54].	69
3.7	Skizze des Kalibrierobjekts mit gesuchten Features für Kamera und LiDAR	71
3.8	Entfernungs- und Intensitätsmessungen von dem Kalibrierobjekt nach Abbildung 3.7.	72
3.9	Kalibrierobjekt in einem Büro [54].	75
3.10	Skalarer Fehler der Kalibrierung gegenüber der Stichprobengröße.	79
3.11	Freizeichnung von belegten Zellen durch zu grobe Auflösung des Belegtheitsgitters.	83
3.12	Beispielszenario für ungewollte Freizeichnung.	85
3.13	Strahlenverlauf innerhalb des Belegtheitsgitters, der zur ungewollten Freizeichnung führt.	89
3.14	Endlicher Automat zur Umsetzung der bedingten Freizeichnung innerhalb eines Strahls R_p . Nach [42].	90
3.15	Ausgangssituation, ohne Berichtigung der ungewollten Freizeichnung.	93
3.16	Ansatz A für das Beispiel einer planaren Bodenplatte	95
3.17	Ansatz B für das Beispiel einer planaren Bodenplatte	96
3.18	Ansatz C für das Beispiel einer planaren Bodenplatte	97
3.19	Gegenüberstellung der Parametersets A(3), B(4) und C(4).	99
3.20	Direktes Einfügen einzelner Scans gegenüber einem kombinierten Einfügen mehrerer Scans.	101
4.1	Approximation einer Objektoberebene mittels einer Lasermessung und eines Kamerabildes.	110
4.2	Idealsituation für die Sichtfelderweiterung mittels dreier Lasermessungen.	111
4.3	Approximation einer Objektoberebene mit n -kollinearen Lasermessungen und einem Kamerabild.	112
4.4	Approximation einer Objektoberebene mit n -linear unabhängigen Lasermessungen und einem Kamerabild.	113

4.5	Differenzierte Kennzeichnung eines Belegtheitsgitters ausgehend von Rohdaten oder einer Regressionsebene.	114
4.6	Blockdiagramm der Sichtfelderweiterung.	117
4.7	Blockdiagramm der Vorverarbeitung der Kamerabilder.	119
4.8	Blockdiagramm der Vorverarbeitung der Laserdaten.	122
4.9	Blockdiagramm der kooperativen Fusion von Kamerabildern und Laserscannerdaten.	123
4.10	Identifikation von Konturpixeln mittels korrespondierenden Lasermessungen.	128
4.11	Erstellung von Segmenten auf Basis von Kantenpixeln.	131
4.12	Berücksichtigung der Laserspotfläche zur Ermittlung von pseudo-Lasermessungen auf der Bildebene.	133
4.13	Identifikation von Konturpixeln mittels Lasermessungen, welche die halbe Winkelauflösung der Kamera aufweisen.	134
4.14	Planare Oberflächenapproximation als Ergebnis der kooperativen Fusion.	137
4.15	Blockdiagramm der konkurrierenden Fusion zur Kombination der vorausgehenden Fusionsergebnisse.	140
4.16	Abweichung planarer Oberflächenapproximation bei nicht planaren Oberflächen.	141
4.17	Signifikanzen bei unterschiedlichen Varianzen von approximierten Objektoberflächen gegenüber deren Rohdaten.	144
4.18	Reduktion der Signifikanz auf Grundlage der Entfernung einer Approximation \tilde{p} zur Lasersicht.	145
4.19	Signifikanz einer Approximation auf Grundlage des entfernungsabhängigen Kameraauflösungsvermögens.	148
4.20	Diskrete Approximation \tilde{p} ohne Sensormodell integriert in OcTree.	149
4.21	Diskrete Messungen \tilde{p} mit Sensormodell in OcTree integriert.	150
4.22	Betrachtung eines Objektes aus mehreren Blickwinkeln.	151
4.23	Integration der diskreten Approximation \tilde{p} von drei Prognosen in den OcTree.	152
4.24	Rekonstruktion der erwarteten Objektoberfläche durch Integration von n Oberflächenprognosen in den OcTree.	153
4.25	Unbemannter Helikopter midiARTIS.	154
4.26	Schematischer Aufbau der ARTIS-Avionik mit relevanten Datenpfaden.	156
4.27	Rohdaten vom monogonen Laserscanner und von der monokularen Kamera.	158
4.28	Vorverarbeitung der Kameradaten.	159
4.29	Segmentierung der Bildebene.	160

4.30	Resultat der kooperativen Fusion.	161
4.31	OcTree, wobei die Färbung der Voxel deren Belegungswahrscheinlichkeit $p(Z)$ widerspiegelt.	162
4.32	Detektionsfehler bei der Bestimmung der Segmentgrenzen, als Resultat einer nicht korrekten Kantendetektion.	163
4.33	OcTree OT_T des Beispielszenarios, erstellt allein mit LiDAR-Daten des monogonen Scanners.	165
4.34	OcTree OT_T des Beispielszenarios erweitert mit Regressionsebene E_G	166
4.35	Verifikation der Regressionsebene E_G gegenüber Laserrohdaten.	166
4.36	Ausgangssituation allein mit Laserdaten.	168
4.37	Inkrementelle Dekomposition der Signaldetektion für den OcTree $OT_{F,1:i_S}$ gegenüber dem Signal OT_S	169
4.38	Detektionsleistung basierend auf der Gegenüberstellung von $OT_{L,1:i_S}$ und OT_T sowie $OT_{F,1:i_S}$ und OT_S	171

Abkürzungsverzeichnis

- ARTIS** Autonomous Rotorcraft Test bed for Intelligent Systems
- CCD** lichtempfindliche elektronische Bauelemente, (engl. Charge-Coupled Device)
- DGPS** Globales Positionssystem mit Differentialsignal, (engl. Differential Global Positioning System)
- FCC** Flugsteuerrechner (engl. Flight Control Computer)
- FIFO** Abarbeitungsprinzip „der Reihe nach“ (engl. First In First Out)
- GCS** Bodenstation (engl. Ground Control Station)
- LiDAR** engl. Light Detection and Ranging
- LRF** engl. Laser Range Finder
- IMU** Inertial-Sensoren-Plattform, (engl. inertial measurement unit)
- OS** Betriebssystem, (engl. Operation System)
- RC** Fernsteuerung, (engl. Remote Control)
- RPAS** Ferngesteuertes Luftfahrzeug, (engl. Remotely Piloted Aircraft Systems)
- UA** unbemanntes Luftfahrzeug, (engl. Unmanned Aircraft)
- VFC** Umgebungsdatenverarbeitungsrechner, (engl. Vision Flight Computer)
- VGA** engl. Video Graphics Array
- WLAN** Kabelloses lokales Netzwerk, (engl. Wireless Lokal Area Network)
- XGA** engl. Extended Graphics Arra

Symbolverzeichnis

Skalare

u, v	diskretisierte Koordinaten auf der Bildebene $u, v \in \mathbb{N}$
w_Z	Durchmesser eines Laserstrahls in der Entfernung Z zur Strahlenquelle
x_n, y_n, z_n	Objektpunktkoordinaten im Bezugssystem n
γ	Divergenzwinkel eines Laserstrahls
Φ, Θ, Ψ	geodätische Lagewinkel des Sensorträgers

Vektoren

\mathbf{p}_n	dreidimensionaler Objektpunkt bezüglich des Koordinatensystems n (siehe Indizes)
$\tilde{\mathbf{p}}$	approximierter dreidimensionaler Objektpunkt bezüglich des Koordinatensystems n (siehe Indizes), dem keine direkte Messung zugrunde liegt.
\mathbf{q}	Bildpunkt
\mathbf{q}°	bezogen auf den Hauptpunkt normierter Bildpunkt
\mathbf{q}'	normierter, von Verzerrung (optischen Aberrationen) bereinigter Bildpunkt
\mathbf{t}_{bt}	Translation eines Koordinatensystems t bezüglich des Koordinatensystems b (siehe Indizes)
Θ_{bt}	Rotation eines Koordinatensystems t bezüglich des Koordinatensystems b (siehe Indizes)

Matrizen

I	Einheitsmatrix
K	Kameramatrix
R_n	Rotationsmatrix einer Rotation um die n -Achse
R_{bt}	Rotationsmatrix, die eine Rotation des Koordinatensystems t bezüglich des Koordinatensystems b beschreibt (siehe Indizes)
bT_t	Homogene Transformationsmatrix, die eine Rotation und Translation des Koordinatensystems t bezüglich des Koordinatensystems b beschreibt (siehe Indizes)

Felder

A	Scan, Kombination aller Messungen p_l , die ein Laserscanner bei einer kompletten Abtastung seines Sichtfeldes erzeugt
B_t	Kamerabild, aufgezeichnet zum Zeitpunkt t

Mengen

C	Punktwolke, Kombination mehrerer dreidimensionaler Messungen p
-----	--

Indizes

c	körperfestes Koordinatensystem der Kamera
b	körperfestes Koordinatensystem des Sensorträgers
o	lokales Weltkoordinatensystem
l	körperfestes Koordinatensystem des Laserscanners

s körperfestes Koordinatensystem eines allgemeinen Sensors,
körperfestes Koordinatensystem des Smartsensors beste-
hend aus Kamera und Laserscanner

Operatoren

$(\dots)^{-1}$ inverse einer quadratischen Matrix

$(\dots)^T$ transponierter Vektor bzw. Matrix

Ebenen

E_A Ebene, welche die Messungen \mathbf{p}_l eines monogonen Scans \mathbf{A}
enthält

E_B Bildebene

$E_{R,C}$ Regressionsebene, auf der Basis der Punktemenge \mathbf{C}

E_O als planar angenommene Oberfläche eines Objektes

E'_O planare Approximation einer Objektoberfläche

\hat{E}_O finites Polygon auf der Approximation E'_O

Kapitel 1

Einleitung

1.1 Motivation

Unbemannte Luftfahrzeuge ermöglichen neue Nutzungskonzepte, welche teilweise außerhalb des klassischen Betätigungsfeldes der bemannten Luftfahrt liegen. Durch den Wegfall des menschlichen Piloten an Bord des Luftfahrzeuges sind kleinere Konstruktionen möglich, welche den Einsatz von Luftfahrzeugen in Umgebungen und Szenarien erlauben, die mit manntragenden Vehikeln nicht erreicht werden können. Der Wegfall des Piloten an Bord des unbemannten Luftfahrzeugs (engl.: Unmanned Aircraft, UA) wirft die Frage auf, wie dessen Fähigkeiten an Bord substituiert werden können. Ein Punkt, dem dabei viel Aufmerksamkeit entgegengebracht wird, ist die Wahrnehmung der Umwelt, welche das UA umgibt, um die Kollision mit möglichen Hindernissen zu vermeiden.

Denkbare Beispiele für zukünftige Einsatzkonzepte von UA sind etwa Individual-Transporter für Konsumgüter wie Pakete 1.1c) [9], Lebensmittel und Medikamente [94], Überwachungsplattformen für Polizei [3] und Rettungskräfte (Abbildung: 1.1b) oder der Einsatz als fliegendes Auge zur Inspektion von technischen Anlagen wie Staudämmen [56] oder Windkraftanlagen (Abbildung: 1.1a) [10].

Viele der möglichen Nutzungskonzepte beziehen sich auf den bodennahen Luftraum zwischen Objekten wie Bäumen und Gebäuden oder gar innerhalb dieser in nächster Nachbarschaft zum Menschen. Diese angestrebte Einsatzumgebung ist dynamisch und verändert sich kontinuierlich. Die Nutzung von UA in diesem Umfeld setzt ein stetiges Monitoring der möglichen Umweltveränderungen voraus, um mögliche Kollisionen des UA mit Objekten wie Mensch, Baum oder Gebäude zu verhindern. Diese Detektionsleistung wird aktuell vom Sicherheitspiloten oder Operateur des UA am Boden geleistet.

Eine automatische Detektion von Hindernissen würde neben der oft vorliegenden Automatisierung der Flugregelung und Pfadplanung weitere Perspektiven für die Nutzung von UA eröffnen.



(a) Inspektion einer Windkraftanlage mit einem ferngesteuerten UA (Remote Piloted Aircraft System RPAS). [48]



(b) Inspektion einer Gefahrenstelle durch die Feuerwehr mit einem UA. Foto: Microdrones



(c) Transport eines Paketes mit einem UA. Foto: Microdrones

Abbildung 1.1: Potenzielle Nutzungsszenarien für UA.

Anwendungen sind etwa Assistenzsysteme, die einen Operator vor der Kollision des UA mit einem Hindernis warnen. Ein mögliches Einsatzbeispiel dafür ist die Inspektion von Windkraftanlagen oder Hochspannungsanlagen. Bei derartigen Inspektionen liegen einerseits größere Entfernungen zwischen UA und Operator vor, die eine Abschätzung von Abständen zwischen UA und zu inspizierendem Objekt erschweren, andererseits sind diese Arbeiten oft monoton, was zu einem Rückgang der Konzentrationsleistung des Operators führt. Beide Faktoren können ursächlich für einen Kontakt des UA mit dem zu begutachtenden Objekt sein, der durch den Einsatz eines Assistenzsystems verhinderbar ist.

Ein weiterer Aspekt ist der automatisierte Betrieb von UA außerhalb der direkten Sicht oder ohne direkte Eingriffsmöglichkeiten des Operators. Dabei würden die von der Hindernisdetektion ermittelten Hindernisse direkt in einer Pfadplanung berücksichtigt. Der Betrieb eines UA außerhalb

der direkten Sicht ist besonders bei Szenarien interessant, die für Leib und Leben des Operateurs eine Gefahr darstellen wie etwa die Kartierung eines kontaminierten Gebiets.

Eine umfassende Detektion der umgebenden Hindernisse wird oftmals durch die beschränkten Fähigkeiten der zur Verfügung stehenden Sensoren verhindert. Die Miniaturisierung von UA verstärkt diese Problematik oftmals, da der Einsatz von klassischen, aus der Luftfahrt bekannten Sensoren aufgrund von Beschränkungen des Bauraums und der möglichen Nutzlast unmöglich ist. Damit ist die Auswahl an möglichen Sensoren weiter eingeschränkt. Um UA für die zukünftigen Einsatzkonzepte vorzubereiten, werden in anderen Forschungsschwerpunkten verwendete Sensoren und Konzepte genutzt und angepasst, um den Kontakt von UA mit Hindernissen zu verhindern.

1.2 Automatische Hindernisdetektion

Die vorliegende Arbeit beschreibt den Aufbau und die Validierung einer berührungslosen Hindernisdetektion für ein kleines UA, im Speziellen einen kleineren unbemannten Hubschrauber¹. Das angestrebte Einsatzgebiet bzw. Szenario für das UA ist der Betrieb außerhalb von Gebäuden innerhalb einer komplexen und zumindest teilweise unbekanntem Hinderniskulisse. Bei dem Betrieb des UA in einer bodennahen Hinderniskulisse wird davon ausgegangen, dass sich der Großteil der in der Kulisse befindlichen Hindernisse nicht kooperativ verhält. Im vorliegenden Fall bedeutet dies, dass die Hindernisse ihre jeweilige Position, Lage oder Größe nicht publizieren. Das gewünschte Szenario erfordert eine kontinuierliche Wahrnehmung der Umwelt während des Fluges und die unverzügliche Bereitstellung der gewonnenen Daten an nachgelagerte Prozesse wie die Pfadplanung bzw. die Kollisionsvermeidung.

UA zeichnen sich durch die Fähigkeit aus, eine dreidimensionale Positionierung im Raum zu ermöglichen und folglich Hindernissen ebenfalls dreidimensional auszuweichen. Um diese Fähigkeit nicht zu beschneiden, bedarf es eines dreidimensionalen Monitorings des Umfelds und einer dreidimensionalen Erfassung der Hindernisse.

Die komplette Wahrnehmung und Verarbeitung erfolgt an Bord des UA, um einen kollisionsfreien Flug ebenfalls dann sicherzustellen, wenn etwa die Datenverbindung zu einer externen Berechnungs- bzw. Kommandoeinheit abbricht. Die Umweltwahrnehmung ist wie Pfadplanung oder Flugregelung ein substanzieller Bestandteil des Betriebs eines UA und ist

¹MTOW \leq 25 kg

keine Nutzlast. Der kollisionsfreie Flug in einem Hindernisszenario ist kein Selbstzweck. Er ist vielmehr Voraussetzung zur Bewältigung von Aufgaben wie Überwachung und Inspektion.

Die Fokussierung auf ein kleines UA und das dargelegte Szenario stellt keine Beschränkung der Allgemeingültigkeit dar. Vielmehr werden die kleinen unbemannten Systeme als Ausgangspunkt für weitere Entwicklungen verstanden. Es wird davon ausgegangen, dass die entwickelte Technik ohne Beschränkung der Allgemeinheit ebenso auf großen unbemannte als auch auf bemannte Systeme übertragen werden kann. Diese Annahme beruht auf der Tatsache, dass in der bemannten Luftfahrt Sensoren² eingesetzt werden, welche Hindernisse in unbekanntem Gelände detektieren. Diese Sensoren sind allerdings sehr komplex, groß und schwer, was einen Einsatz auf einem kleinen UA verhindert und gegenwärtig eine umgekehrte Entwicklungsrichtung der Generalisierung von großen zu kleinen Flugsystemen erschwert.

1.2.1 Beschaffung von Umgebungsinformationen

Zur Hinderniswahrnehmung werden unter anderem Sensoren wie bspw. Kameras, Sonar, Radar oder Laserscanner (engl.: Light Detection and Ranging, LiDAR) eingesetzt.

Kameras sind passive Sensoren und nutzen die in der Umwelt bereitgestellte Energie, um Messungen vorzunehmen. Die Möglichkeit, ohne eine eigene Energiequelle auszukommen, ermöglicht im Allgemeinen kleinere und leichtere Bauformen als bei aktiven Sensoren. Allerdings sind passive Sensoren auf die Qualität der von der Umwelt bereitgestellten Energie angewiesen. In stabilen, bekannten und reproduzierbaren Umweltbedingungen ist der Einsatz von passiven Sensoren weitestgehend unproblematisch, da die Sensoren auf die vorherrschende Situation eingestellt werden können und folglich gute Ergebnisse erzielen. Dynamische bzw. schlecht reproduzierbare Umweltbedingungen, wie sie außerhalb von Gebäuden auftreten, erschweren den Einsatz von passiven Sensoren, was etwa bei Kameras zu Über- bzw. Unterbelichtungen führen kann.

Aktive Sensoren sind weniger anfällig für die dynamischen Umweltverhältnisse außerhalb von Gebäuden. Kleine zweidimensionale Systeme, welche von kleinen UA getragen werden können, geben jedoch nur ein sehr begrenztes Abbild der Umwelt wieder. So decken die Sichtfelder eines zweidimensionalen LiDAR oft nicht einmal die Abmessungen des UA ab, was zu Kollisionen mit Hindernissen führen kann, die außerhalb der Sensorsicht, aber innerhalb des Raumes liegen, der von dem UA beansprucht

²z.B. Helicopter Laser Radar (kurz HELLAS) der Firma Cassidian

wird. Umfangreichere Sichtfelder benötigen aufwendigere Systeme, deren zusätzliches Gewicht den Einsatz auf einem kleinen UA verhindern. Sowohl Kameras als auch kleine, leichte, aktive Sensoren eignen sich aufgrund ihrer Eigenschaften nur bedingt als alleinige Datenquellen für eine dreidimensionale Hindernisdetektion. Aus diesem Grund sollen die Daten beider Sensorkonzepte in der vorgestellten Arbeit kombiniert werden, um eine umfassendere Sicht auf die Umwelt zu erreichen. Dafür wird ein zweistufiges Konzept vorgeschlagen. Im ersten Schritt, einer kooperativen Fusion³, werden dreidimensionale Approximationen generiert, um einen Sichtkegel auf die Umwelt zu erhalten, der über das Sichtfeld des Lasers hinausgeht und dem Sichtfeld der Kamera maximal entspricht. Im zweiten Schritt, einer konkurrierenden Fusion⁴, werden die zuvor generierten Approximationen konkurrierend in ein dreidimensionales Weltmodell integriert. Dies geschieht einerseits, um Abweichungen zwischen Approximation und realer Umwelt zu reduzieren, welche auf Annahmen der kooperativen Fusion fußen. Andererseits ermöglicht die Fusion mehrerer Datensätze über die Zeit eine weitere verdichtete Anreicherung der Sensordaten, um eine globale Hindernisrepräsentation zu erhalten. Abbildung 1.2 visualisiert dieses Konzept in einem Blockdiagramm.

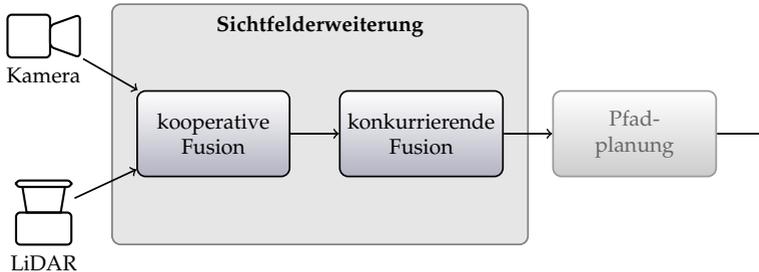


Abbildung 1.2: Blockdiagramm der Sichtfelderweiterung.

³ Kombination von Daten mehrerer Sensoren, wobei erst die Kombination die Bereitstellung der gewünschten Information ermöglicht. Die Daten der Einzelsensoren enthalten jeweils nicht den gewünschten Informationsumfang.

⁴ Kombination von gleichartigen Daten, welche von unabhängigen Sensoren und von einem Wahrnehmungsbereich stammen. Die Kombination erfolgt oft gewichtet, um die Genauigkeit des Gesamtsystems zu erhöhen, oder um Unterschiede zwischen den Daten zu identifizieren. Das Ziel ist eine Steigerung der Aussagewahrscheinlichkeit und eine erhöhte Fehlertoleranz für den gemeinsamen Wahrnehmungsbereich.

1.2.2 Kooperative Fusion

Im hier dargelegten Ansatz werden die Daten einer monokularen Kamera und ein zweidimensionales LiDAR⁵ kombiniert, um die dreidimensionale Form von Hindernisoberflächen zu approximieren.

Monokulare Kameras stellen ausschließlich relativ skalierbare bzw. skalierungsinvariante Daten bereit, aus denen direkt auf Basis eines Einzelbildes keine dreidimensionalen Informationen abgeleitet werden können. Lasertfernungsmesser erlauben die direkte Bereitstellung dreidimensionaler Informationen, allerdings sind diese bei dem verwendeten zweidimensionalen LiDAR auf eine Zeile begrenzt. Was sich unter- bzw. oberhalb dieser Zeile befindet, ist unbekannt.

Innerhalb der Fusion werden auf Grundlage der Entfernungsmessungen des LiDAR dreidimensionale Aussagen für ausgewählte Bereiche der Kamerasicht getroffen, für welche direkt keine Lasermessungen vorliegen. Dies ermöglicht eine Steigerung der dreidimensionalen Umweltwahrnehmung auf die maximale Größe des Kamerasichtfeldes. Die Erweiterung der dreidimensionalen Sicht auf Bereiche der Kamerasicht außerhalb der Lasersicht basiert auf der Approximation von Objektoberflächen, welche teilweise vom LiDAR wie auch von der Kamera betrachtet werden.

Einen einfachen Aufbau dieses Prinzips zeigt Abbildung 1.3. Auf Basis von drei unabhängigen Lasermessungen, welche für eine Objektoberfläche vorliegen, und der getroffenen Annahme, dass die betrachtete Oberfläche planar ist, wird eine Regressionsebene der betrachteten Oberfläche erstellt. Die bestimmte Regressionsebene ist der Ausgangspunkt, um für ausgewählte Pixel der Bildebene ebenfalls Entfernungsmessungen zu bestimmen.

Die Kameramessungen sind skalierungsinvariant⁶ und erlauben ausschließlich die Berechnung eines Strahls, auf dem eine Messung liegt. Die Lage der Messung auf dem Strahl hingegen ist unbekannt. Die Ermittlung des Schnittpunkts zwischen einem Lichtstrahl, der von der Kamera eingefangen wird, und der Regressionsebene stellt somit eine approximierte Entfernung zwischen Kamera und angenommener Objektoberfläche bereit. Somit könnte für ausgewählte Pixel der Kamera neben der Beleuchtungs-

⁵Als zweidimensionales LiDAR wird in dem hier präsentierten Fall eine monogoner Laserscanner verstanden, welcher pro Messung ein Winkel ϕ und eine Entfernung z bzgl. des Sensorkoordinatensystems bereitstellt.

⁶Merkmale bzw. Messungen sind skalierungsinvariant, wenn keine unabhängige Aussage zur Größe und Entfernung eines betrachteten Objekts, ohne die Verwendung absoluter Aussagen, getroffen werden kann. Objekte mit identischem Aussehen, unterschiedlicher Größe und in unterschiedlicher Entfernungen zur Kamera positioniert, können auf der Bildebene identische Merkmale generieren.

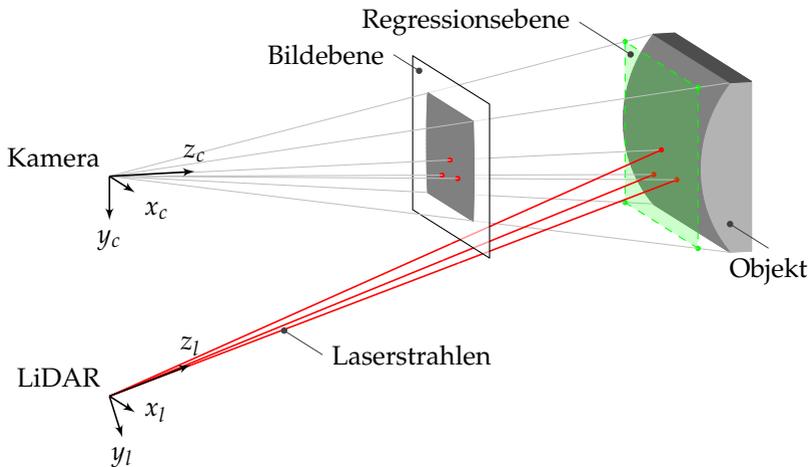


Abbildung 1.3: Approximation einer Objektoberflächenebene mittels unabhängiger Lasermessungen und Kamerabildern.

formation auch eine Entfernungsschätzung bestimmt werden.

Die Zuordnung, ob ein Pixel einer Objektoberfläche zugeordnet werden kann, für welche eine Regressionsebene vorliegt, erfolgt über eine Segmentierung der Bildebene auf Basis der Intensitäten. Die Segmentierung beruht auf der Annahme, dass zusammenhängende Objektoberflächen auf der Bildebene ebenfalls als zusammenhängende Segmente mit einer homogenen Eigenschaft abgebildet werden. Für jeden Pixel eines Segmentes, für welches eine Regressionsebene bestimmt werden kann, kann folglich ein Schnittpunkt bzw. die Entfernung zu diesem berechnet werden.

Je nach Gestalt der Hindernisstruktur, die von Kamera und LiDAR betrachtet wird, können mit diesem Vorgehen dreidimensionale Entfernungsapproximationen für ausgewählte Pixel des Kamerabildes erstellt werden, welche außerhalb der LiDAR-Sicht liegen.

1.2.3 Konkurrierende Fusion

Die auf Basis der drei Lasermessungen (Abbildung: 1.3) approximierte planare Objektoberfläche ist eine Regression und spiegelt die wahre Objektoberfläche nur in Bereichen genähert wieder, in denen auch Lasermessungen vorliegen. In Regionen der betrachteten Oberfläche, welche außerhalb der Lasersicht liegen, können keine Aussagen über die Übereinstimmung von Regression und wahrer Oberfläche gemacht werden, sodass dort Ab-

weichungen vorliegen können. Abbildung 1.3 skizziert diesen Zustand anhand einer gebogenen Objektoberfläche, welche durch eine Ebene approximiert wird. Die Approximation repräsentiert die Objektoberfläche in einer Region nur hinreichend, in der auch Lasermessungen vorliegen.

Aufgrund der Vielzahl von Oberflächenformen ist es nicht praktikabel, für jede Form eine unabhängige Approximation bereitzustellen. Die Annahme einer planaren Objektoberfläche soll beibehalten werden. Allerdings sollten die möglichen Fehler auch nicht vernachlässigt werden.

Es wird die Annahme aufgestellt, dass bei einem Flug eines UA durch eine Hinderniskulisse Hindernisse über die Zeit aus unterschiedlichen Positionen und Blickwinkeln mehrfach betrachtet werden. Dies ermöglicht das Erstellen mehrerer planarer Oberflächenapproximationen, welche sich teilweise bestätigen, aber auch widersprechen können. Eine konkurrierende Kombination der sich bestätigenden und widersprechenden Informationen würde damit die Herausbildung einer realistischeren Approximation der Hindernisoberfläche ermöglichen, als es die alleinige Nutzung der ebenen Approximation vermochte.

Für die konkurrierende Fusion werden die einzelnen Entfernungsapproximationen in eine räumliche Repräsentation übertragen, welche nicht Auskunft über die Oberflächenform, sondern über die räumliche Belegung eines Hindernisses gibt. Die räumliche Repräsentation wird durch ein dreidimensionales, metrisches Gitter aufgebaut, welches den Raum, der das UA umgibt, je nach Auflösung hinreichend fein diskretisiert. Die Elemente des Gitters geben probabilistisch Auskunft darüber, ob der Raum, den sie beschreiben, frei von Hindernissen oder belegt ist.

Die Integration der unabhängigen Approximation in das Gitter erfolgt gewichtet, was eine kombinierte Entscheidung über die jeweilige Belegung einer Zelle erlaubt. Die Gewichtung jeder Approximation wird mittels eines Sensormodells umgesetzt, welches unter anderem die Lage der Approximation bezüglich der rohen Lasermessungen berücksichtigt und somit jeder Approximation einen Vertrauenswert zuspricht.

Das Resultat der konkurrierenden Fusion ist eine Hinderniskarte, welche Hindernisoberflächen nicht als Flächen, sondern als Belegungen des Raumes beschreibt, welche von einem UA nicht durchdrungen werden können. Die Kombination gestattet damit je nach Auflösung des Gitters eine realistischere Abbildung der Hindernisoberflächen, als es die planare Approximation der kooperativen Fusion erlaubt. Die Karte erlaubt darüber hinaus den Aufbau größerer Hindernisstrukturen, welche weit über die Sichtfelder der Einzelsensoren hinausgehen, und die Behandlung von dynamischen Hindernissen.

Das dargelegte Konzept lässt sich in zwei Kernthesen zusammenfassen:

1. Aufbauend auf einer kooperativen Fusion von monokularen Kamera- und monogonen⁷ LiDAR-Daten lässt sich eine Sichtfelderweiterung für das LiDAR generieren, welche es ermöglicht, dreidimensionale Approximationen von Hindernisoberflächen außerhalb der Lasersicht zu generieren, wobei die Hindernisoberflächen, welche durch die Approximationen beschrieben werden, teilweise in der Kamera- als auch Lasersicht vorliegen.
2. Die erstellten dreidimensionalen Approximationen lassen sich unter Verwendung eines Sensormodells in einer metrischen, kubischen Umweltrepräsentation seriell konkurrierend kombinieren, um eine hinreichend zutreffende Hindernisform herauszuarbeiten und aufbauend den Einsatz in einer globalen Hinderniskartierung zu ermöglichen.

1.3 Aufbau der Arbeit

Aufbauend auf der dargelegten Motivation und dem vorgestellten Konzept in Kapitel 1 „Einleitung“ schließt sich eine Darstellung der benötigten Grundlagen in Kapitel 2 an. Es erfolgt eine Vorstellung der geometrischen Grundlagen, wie verwendete Koordinatensysteme und Transformationen sowie der verwendeten Sensoren, Kamera und LiDAR. Darüber hinaus wird eine Einführung in die Segmentierung der Bildebene und die Repräsentation der Umwelt mittels eines OcTree gegeben.

Den Kern der Arbeit bilden die beiden Kapitel 3 „Basisqualifikationen“ und 4 „Sichtfelderweiterungen“.

In den Basisqualifikationen werden die zwei Punkte „Bestimmung der extrinsischen Parameter zwischen Kamera und LiDAR“ und „Transformation der kontinuierlichen Umwelt in eine diskrete Repräsentation“ behandelt. Beide Punkte sind Grundbausteine, ohne welche die eigentliche Sichtfelderweiterung nicht realisiert werden kann. Das Kapitel 4 beschreibt den bisherigen Stand der Technik, das Konzept sowie die daraus erarbeiteten Algorithmen, die für Sichtfelderweiterungen erforderlich sind. Die Aufarbeitung der Sensordaten und die Kombination werden in den drei Abschnitten „Vorverarbeitung der Sensordaten“, „kooperative Fusion“ und

⁷ Ein monogones Objekt ist u.a. ein zweidimensionales geometrisches Objekt, das einen einzigen Eckpunkt besitzt. Dieser Eckpunkt ist Start- und Endpunkt einer Kante. Die Kante spannt eine Fläche auf, die komplett von der Kante begrenzt ist. Übertragen auf ein LiDAR ist der Eckpunkt Quelle und Senke des Sensors, und die Messungen liegen in der aufgespannten Fläche.

„konkurrierende Fusion“ aufgeteilt.

Nach der Vorstellung der verwendeten Algorithmen erfolgt eine Validierung der Sichtfelderweiterung anhand eines Beispielszenarios.

Die Ergebnisse im Hinblick auf die Forschungsfragen der Arbeit sowie ein Ausblick auf zukünftige Arbeiten werden in Kapitel 5 gegeben.

Kapitel 2

Grundlagen

Dieses Kapitel bietet einen Überblick zu den Themen Sensordatenfusion, Sensoren der Umweltwahrnehmung, Segmentierung im Bereich der Bildverarbeitung sowie Hindernisrepräsentationen, welche zum Aufbau der angestrebten Sichtfelderweiterung benötigt werden. Darüber hinaus werden einige Koordinatensysteme, Transformationen und Konventionen vorgestellt, auf welche im Laufe der Arbeit wiederholt Bezug genommen wird. Für die Thematik der Sensordatenfusion wird speziell auf Zielstellung, Eigenschaften und Architektur der einzelnen Kombinationen eingegangen. Nach einem Überblick über mögliche Sensoren zur Umweltwahrnehmung, wird speziell auf die beiden Sensoren monokulare Kamera und monogones LiDAR eingegangen. Neben den jeweiligen Eigenschaften stehen Fehlerquellen im Fokus.

Die monokulare Kamera bildet neben dem LiDAR eine Datenquelle der vorliegenden Arbeit. In dem Abschnitt Segmentierung wird ein grundlegender Überblick zur Segmentierung auf der Bildebene gegeben und der Ansatz der Kantendetektion näher beleuchtet.

Der Abschnitt Hindernisrepräsentationen gibt einen Überblick möglicher Ansätze. Besonderes Augenmerk liegt auf der Hindernisrepräsentation mittels eines dreidimensionalen Belegtheitsgitters auf Basis eines OcTrees.

2.1 Sensordatenfusion

Methoden zur Kombination von Sensordaten oder zur Verknüpfung von korrespondierenden Informationen, um eine Datenbasis bzw. Schlussfolgerungen zu generieren, welche zu Informationen führen, die ohne die Zusammenführung der Daten nicht bereit gestanden hätten, werden im Allgemeinen als Sensordatenfusion bezeichnet.

Die Quellen der verwendeten Daten müssen nicht zwangsläufig mehrere unterschiedliche unabhängige Sensoren sein. Auch über den zeitlichen Verlauf generierte Daten eines Sensors können für eine präzisierte Gesamt-

aussage kombiniert werden.

Das Vorgehen, verschiedene Sensordaten zu kombinieren, um eine Situation besser klären zu können, ist keine rein technische Lösung. In der Biologie sind viele Beispiele bekannt, in denen Lebewesen gelernt haben, mehrere ihrer Sinnesorgane einzusetzen, um eine Situation aufzuklären und damit ihr Überleben zu sichern. In Situationen, in denen bspw. die visuelle Detektion von Feinden oder Beute durch Vegetation, Dunkelheit oder biologische Eigenheiten (Sichtfeld, Farbsehen, u.Ä.) begrenzt ist, werden weitere Sinneseindrücke wie Gehör-, Tast- oder Geruchssinn mit der visuellen Wahrnehmung oder auch untereinander kombiniert. Die durch die Fusion generierte Informationsbasis erlaubt möglicherweise eine frühere Flucht oder einen Jagderfolg, welche bei der alleinigen Verwendung der visuellen Sinnesorgane nicht möglich gewesen wären [38].

Die Fusion von Daten von mehreren Sensoren hat mehrere Vorteile gegenüber der alleinigen Nutzung eines Sensors. Geht man von biologisch inspirierten Beispielen aus, steigert die Kombination mehrerer Sensoren mit unterschiedlichen Messprinzipien (visuell, akustisch, haptisch oder olfaktorisch) die Detektionswahrscheinlichkeit von Feinden oder Beute. Die Einschränkung eines Sensors durch bspw. gegebene Umweltbedingungen führt bei einer Fusion von mehreren Sinneseindrücken nicht notwendigerweise zum „Übersehen“ eines Ereignisses. Das Ereignis wird trotz der Einschränkung weiterhin vom Gesamtsystem bzw. Individuum registriert.

Ein identisches Bild zeigt sich bei der Zuverlässigkeit eines Gesamtsystems, das durch die Verwendung mehrerer unabhängiger Sensoren im Allgemeinen steigt. Beim Ausfall eines Sensors kann dessen Aufgabe möglicherweise durch einen weiteren Sensor übernommen werden, und es kommt nicht zum Ausfall des Gesamtsystems.

Durch die Verwendung mehrerer Sensoren, die dasselbe Sichtfeld abdecken, besteht die Möglichkeit, Mehrdeutigkeiten auszuschließen.

Sichtfelder von Sensoren sind aufgrund konstruktiver Eigenheiten sowohl in der Biologie als auch in der Technik begrenzt. Die Kombination mehrerer Sensoren, deren Sicht nicht identisch orientiert ist, ermöglicht es, das Sichtfeld des Gesamtsystems zu vergrößern.

Zusammengefasst werden die Vorteile der Datenfusion von Bellot [12] und Hall [38] in fünf Punkten:

Repräsentanz: Die Resultate der Fusion besitzen eine höhere Abstraktionsebene oder Granularität als die einzelnen Datenquellen.

Bestimmtheit: Durch die Fusion V_F wird eine Verstärkung der Signifikanz oder Glaubwürdigkeit $p(V)$ eines Einzelsensors V mit $p(V_F) > p(V)$ erwartet.

Genauigkeit: Die Standardabweichung der Daten ist nach der Fusion kleiner als für die eingehenden Einzelsensoren.

Vollständigkeit: Jeder neue Informationszusatz trägt zur Komplettierung der Umweltwahrnehmung bei.

Redundanz: Erhöhung der Ausfallsicherheit durch Substitution von Daten bei Verlust einer Fusionskomponente.

Der Einsatz mehrerer Sensoren und die Kombination der so gewonnenen Daten führt allerdings zu den Nachteilen einer gesteigerten Komplexität beim Transport, Verknüpfen und Verwenden der Daten gegenüber der alleinigen Verwendung einer Datenquelle. Ein komplexeres System erhöht wiederum die Wahrscheinlichkeit von Fehlern bei der Planung und Umsetzung der Fusion. Fehlerquellen sind unter anderem:

Nicht korrespondierende Daten: Ein umfassender Situationsüberblick erfordert möglicherweise die Integration von Sensoren bzw. Datenquellen, die hinsichtlich Datenstruktur und/oder -inhalt unterschiedlich sind. Um die Daten in ein semantisches Level zu überführen, auf dem sie kombinierbar werden, sind Vorverarbeitungen notwendig.

Alter der Daten: Die Aktualisierungsrate der bereitgestellten Daten ist in der Regel für jeden Sensor unterschiedlich. Dies führt dazu, dass ein Multi-Sensor-Fusionsaufbau die Verarbeitung von Informationen unterschiedlichen Alters ermöglicht muss.

In der Literatur haben sich drei unterschiedliche Kriterien zu Kategorisierung von Sensordatenfusionen durchgesetzt. Eine Einordnung nach der Funktion der Fusion wird von Brooks und Iyengar [18] präsentiert. Klein [52] trennt die Ansätze basierend auf ihrer Architektur. Hall und Llinas [38] präferieren eine Unterscheidung zwischen den Ebenen innerhalb des Datenflusses.

Fusionsfunktion: Sensordatenfusionen lassen sich unter anderem wie von Brooks und Iyengar [18] präsentiert nach ihrer Funktion kategorisieren. Die Funktion einer Datenfusion wird von Brooks und Iyengar zwischen einer konkurrierenden, komplementären bzw. kooperativen Kombination der Daten unterschieden.

Bei einer konkurrierenden Fusion liefern unabhängige Sensoren gleichartige Daten von einem Wahrnehmungsbereich, wobei die Sensoren nicht das gleiche Messverfahren verwenden müssen. Die generierten Daten werden

oft gewichtet kombiniert, um die Genauigkeit des Gesamtsystems zu erhöhen oder um Fehler bzw. Unterschiede zwischen den Daten zu identifizieren [46]. Das Ziel ist, eine Steigerung der Aussagewahrscheinlichkeit und eine erhöhte Fehlertoleranz für den gemeinsamen Wahrnehmungsbereich zu erreichen.

Die komplementäre Fusion hat das Ziel, die Messungen unabhängiger Sensoren, welche unterschiedliche Wahrnehmungsfelder, Messphänomene oder Messzeitpunkte bereitstellen, zu kombinieren, um die Vollständigkeit der Daten für das Gesamtsystem zu erhöhen.

In der Praxis können die einzelnen Sensoren nicht immer die gewünschten Daten bereitstellen. Erst durch eine Kombination der Daten von mehreren Sensoren ist die Bereitstellung der gewünschten Information möglich. Die Funktionalität dieser Fusion wird als „kooperierend“ bezeichnet.

Fusionsarchitektur: Die Differenzierung einer Fusionsarchitektur erfolgt zwischen Sensor-Level-, Central-Level- und Hybrid-Fusion und ist bedingt durch die Position der Fusion in der Verarbeitungskette der Daten. Die Sensor-Level-Fusion Architektur verwendet für jeden Sensor eine separate Signalvorverarbeitungsstufe, bevor die Daten in dem Fusionsmodul kombiniert werden. In der Signalverarbeitungsstufe werden die Daten des jeweiligen Sensors separat prozessiert. Ein derartiges Vorgehen ist bei sich stark unterscheidenden Sensordaten zu empfehlen. Die frühzeitige Signalverarbeitung führt frühzeitig zu einer Datenreduzierung, was einen reduzierten Datentransfer ermöglicht. Die Ergebnisse der Signalverarbeitung sind je nach Detaillierungsgrad Feature oder Entscheidungen, welche anschließend fusioniert werden.

Die Central-Level-Fusion oder Pixel-Level-Fusion verwendet ausschließlich rudimentär vorverarbeitete Daten. Die Vorverarbeitung beschränkt sich auf Aktionen wie Filtern oder Beleuchtungsschätzung, die vom Sensor automatisch durchgeführt werden und keine separate Signalverarbeitung benötigen. Die nicht vorhandene bzw. geringe Datenreduzierung führt generell zu einer höheren Komplexität beim Datentransport sowie der Prozessierung im Fusionsmodul gegenüber der Sensor-Level-Fusion.

Die Bezeichnung „hybride Fusion“ wird für Mischformen aus Sensor-Level- und Central-Level-Fusion verwendet. Der Ansatz fusioniert sowohl simultan vorverarbeitete Daten als auch Daten, welche direkt vom Sensor bereitgestellt werden. Dies führt im Allgemeinen zu einer komplexeren Architektur als die Verwendung einer der Ausgangsansätze.

Fusionslevel: Eine Einordnung nach dem Level bzw. der Ebene erfolgt nach der Abstraktion der zu fusionierenden Daten. Unterschieden wird

in Data-Level, Feature-Level und Entscheidungs-Level. Beim Data-Level werden die Rohdaten mit einer minimalen Vorverarbeitung verwendet. Diese Kategorie ist der dem Central-Level bei der Kategorisierung nach der Fusionsarchitektur gleichzusetzen.

Eine Klassifizierung nach Feature-Level bzw. Entscheidungs-Level ist schwierig, da der Übergang zwischen Feature und Entscheidung fließend ist und die Generierung eines Features ebenfalls auf Entscheidungen basiert, welche es definieren. Beide Kategorien sind mit der Sensor-Level-Fusion Architektur vergleichbar, da beide Ansätze eine umfangreiche Vorverarbeitung voraussetzen, um die zu fusionierenden Daten (Feature und Entscheidungen) zu generieren.

Ebenso wie bei der Fusionsarchitektur existieren bei der Kategorisierung nach dem Fusionslevel auch Hybride.

2.2 Geometrische Konventionen und Grundlagen

2.2.1 Verwendete Koordinatensysteme

Als lokales Weltkoordinatensystem o dient ein rechtshändiges, kartesisches Koordinatensystem, dessen Ursprung P_o auf der Ellipsoidoberfläche des geodätischen Koordinatensystems e vorab definiert wurde. Das geodätische Koordinatensystem e ist nach WGS84 definiert. Orientiert ist das System o in der Weise, dass die x -Achse tangential in Richtung Norden, die y -Achse orthogonal zur x -Achse, tangential Richtung Osten und die z -Achse orthogonal zur aufgespannten x - y -Ebene in das Innere des Ellipsoiden zeigt. Für das System o wird zur Vereinfachung ein ebenes Erdmodell angenommen. Diese Vereinfachung kann nur für einen kleinen Aktionsradius um den Initialisierungspunkt P_o verwendet werden, große horizontale Aktionsradien um den Punkt P_o führen, basierend auf der Erdkrümmung, zu nicht vernachlässigbaren Fehlern.

Das körperfeste Koordinatensystem des Luftfahrzeuges b ist in dessen Massenschwerpunkt $S = P_b$ verortet. Die Orientierung erfolgt mit der x -Achse entlang der Flugzeuglängsachse, die y -Achse orthogonal dazu durch den rechten Flügel und die z -Achse orthogonal zur x - y -Ebene durch den Boden des Luftfahrzeuges.

Die Sensorkoordinatensysteme s_n sind bezüglich des Koordinatensystems b des Luftfahrzeuges oder weiterer Sensorsysteme s_n beschrieben. Die Sensorsysteme sind dementsprechend orientiert, dass entlang der Hauptsichtachse des Sensors die z -Achse verläuft. Die x -Achse liegt orthogonal zur

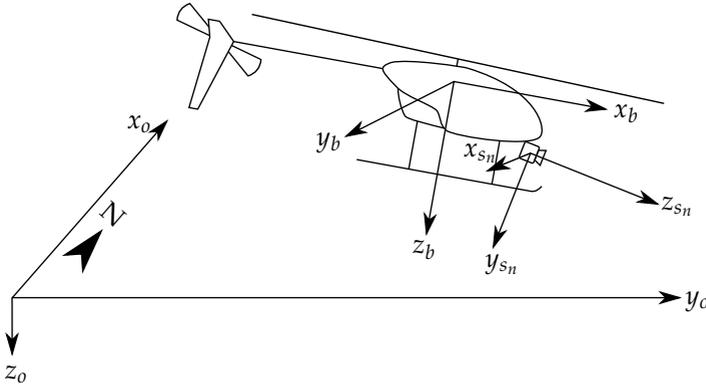


Abbildung 2.1: lokales Welt-, körperfestes Träger- und Sensorkoordinatensystem

z -Achse, um die horizontale Auslenkung der Messwerte gegenüber der Hauptachse zu beschreiben. Die y -Achse liegt orthogonal zur x - z -Ebene und ermöglicht die Beschreibung der vertikalen Auslenkung der Messungen. In dem präsentierten Fall werden zur einfacheren Verdeutlichung das Kamerakoordinatensystem mit c und das Laserkoordinatensystem mit l bezeichnet. Das c -System ist gegenüber der Flugplattform b und das System l gegenüber dem c -System beschrieben.

2.2.2 Koordinatentransformation

Überblick

Die Bereitstellung der jeweiligen Daten und ihre Verwendung findet nicht notwendigerweise in ein und demselben Koordinatensystem statt. Aus diesem Grund sind die Daten zwischen den einzelnen Systemen zu transformieren. Beispielhaft soll eine Transformation einer dreidimensionalen Punktmessung \mathbf{p} zwischen Sensorsystem und körperfestem Koordinatensystem des Sensorträgers verdeutlicht werden. Die Transformation des s -Systems bezüglich des b -Systems wird mittels des Orientierungsvektors $\Phi_{bs}(\Phi, \Theta, \Psi)^T$ und des Translationsvektors $\mathbf{t}_{bs}(x, y, z)^T$ beschrieben. Mittels einer Rotationsmatrix \mathbf{R}_{bs} , welche äquivalent zu dem Rotationsvektor Φ_{bs} ist, lässt sich die Relation

$$\mathbf{p}_s = \mathbf{R}_{bs} \mathbf{p}_b + \mathbf{t}_{bs} \quad (2.1)$$

herleiten. Die Rotationsmatrix R ist eine quadratische Matrix, wobei $\mathbf{R} \in \mathfrak{R}^{n \times n}$, $\mathbf{R}^T \mathbf{R} = \mathbf{R} \mathbf{R}^T = \mathbf{I}_n$ und damit $\mathbf{R}^{-1} = \mathbf{R}^T$ gilt. Für die entgegengesetzte Transformation gilt damit folglich

$$\mathbf{p}_b = \mathbf{R}_{bs}^{-1}(\mathbf{p}_s - \mathbf{t}_{bs}). \quad (2.2)$$

Die Rotationsmatrix R_{bs} kann in drei Einzelrotationen

$$\begin{aligned} \mathbf{R}_x(\Phi) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\Phi) & \sin(\Phi) \\ 0 & -\sin(\Phi) & \cos(\Phi) \end{pmatrix}, \\ \mathbf{R}_y(\Theta) &= \begin{pmatrix} \cos(\Theta) & 0 & -\sin(\Theta) \\ 0 & 1 & 0 \\ \sin(\Theta) & 0 & \cos(\Theta) \end{pmatrix} \text{ und} \\ \mathbf{R}_z(\Psi) &= \begin{pmatrix} \cos(\Psi) & \sin(\Psi) & 0 \\ -\sin(\Psi) & \cos(\Psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \quad (2.3)$$

um die jeweiligen Achsen x , y und z des kartesischen Koordinatensystems zerlegt werden. Die Kombination der Einzelrotationen erfolgt nach der (z, y', x'') -Konvention [32, S.164]

$$\mathbf{R}_{bs} = \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x, \quad (2.4)$$

wobei die Drehachsen mitgedreht werden.

Eine in der Robotik verbreitete Darstellung einer Transformation, in der \mathbf{R}_{bs} und \mathbf{t}_{bs} kombiniert werden, ist die homogene Transformationsmatrix. Diese ermöglicht eine ausschließliche Nutzung von Matrizenmultiplikationen ohne eine zusätzliche Anwendung einer Addition, wie es bei 2.1 notwendig ist.

$${}^b \mathbf{T}_s = \begin{bmatrix} \begin{bmatrix} \mathbf{R}_{bs} \\ \mathbf{t}_{bs} \end{bmatrix} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \end{bmatrix} \quad (2.5)$$

Die alleinige Nutzung der Matrizenmultiplikation erlaubt es, eine assoziative Kombination mehrerer aufeinanderfolgender Transformationen zu

einer einzelnen Gesamttransformation zusammenzufassen. Bei der Kombination ist aufgrund der fehlenden Kommutativität die Reihenfolge der Manipulationen zu beachten. Die Notation ${}^s\mathbf{T}_b$ verdeutlicht eine Transformation einer dreidimensionalen Position im System s bezüglich des Systems b , wobei

$${}^s\mathbf{T}_b = {}^b\mathbf{T}_s^{-1} \quad (2.6)$$

entspricht. Die Transformation 2.1 kann mit der homogenen Transformationsmatrix ${}^b\mathbf{T}_s$ wie folgt

$$\begin{bmatrix} \left[\begin{array}{c} \mathbf{p}_s \\ 1 \end{array} \right] \end{bmatrix} = {}^b\mathbf{T}_s \begin{bmatrix} \left[\begin{array}{c} \mathbf{p}_b \\ 1 \end{array} \right] \end{bmatrix} \quad (2.7)$$

umgesetzt werden.

2.2.3 Verwendete Koordinatentransformation

Eine Georeferenzierung der durch die Sensoren bereitgestellten Messungen erfordert eine Transformation der Sensordaten in das lokale Weltkoordinatensystem. Die Transformation wird in mehrere Einzelkomponenten zerlegt. Abbildung 2.2 zeigt die verwendeten Koordinatensysteme sowie die benötigten Transformationen, um dreidimensionale Messpunkte zwischen ihnen zu übertragen. Die Transformation eines Bildpunktes \mathbf{p}_c aus dem Kamerakoordinatensystem c in das System o erfolgt über das Trägerkoordinatensystem. Die Pose ${}^o\mathbf{T}_b$ des Trägers ist bezüglich des Systems o gegeben. Die Pose ${}^b\mathbf{T}_c$ der Kamera ist wiederum gegenüber dem Träger definiert. Aufbauend darauf ergibt sich für die Transformation

$${}^o\mathbf{T}_c = {}^o\mathbf{T}_b \cdot {}^b\mathbf{T}_c. \quad (2.8)$$

Die Invertierung von ${}^o\mathbf{T}_c$,

$$({}^o\mathbf{T}_c)^{-1} = ({}^b\mathbf{T}_c)^{-1} \cdot ({}^o\mathbf{T}_b)^{-1} \quad (2.9)$$

erlaubt die anschließende Transformation von \mathbf{p}_c nach

$$\mathbf{p}_o = ({}^o\mathbf{T}_c)^{-1} \cdot \mathbf{p}_c. \quad (2.10)$$

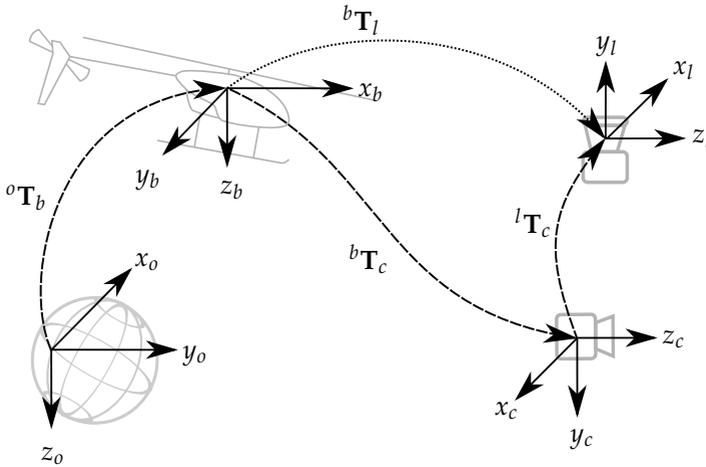


Abbildung 2.2: Transformationen zwischen Welt-, Träger-, Kamera- und Laserkoordinatensystem

Die Transformation des Laserscanners gegenüber dem Sensorträger ist nicht direkt gegeben. Die Pose des Laserscanners ist gegenüber dem Kamerakoordinatensystem definiert. Die Transformation ${}^b\mathbf{T}_o$ zwischen Laser- und Trägerkoordinatensystem ist wie folgt definiert

$${}^b\mathbf{T}_l = ({}^c\mathbf{T}_l)^{-1} \cdot ({}^b\mathbf{T}_c)^{-1}. \quad (2.11)$$

Eine Transformation eines Laserpunktes \mathbf{p}_l in das lokale Weltkoordinatensystem ergibt sich mittels

$$\begin{aligned} ({}^o\mathbf{T}_l)^{-1} &= ({}^b\mathbf{T}_l)^{-1} \cdot ({}^o\mathbf{T}_b)^{-1} \\ \mathbf{p}_o &= ({}^o\mathbf{T}_l)^{-1} \cdot \mathbf{p}_l. \end{aligned} \quad (2.12)$$

Voraussetzung der Transformationen 2.10 und 2.12 sind die Kenntnis der Pose des Trägers im lokalen Weltkoordinatensystem, der Pose der Kamera am Träger sowie der Pose des Lasers gegenüber der Kamera. Auf das Vorgehen, wie die Pose des Trägers im lokalen Weltkoordinatensystem und die Ausrichtung der Kamera am Träger bestimmt werden, wird im anschließenden Abschnitt eingegangen. Die Bestimmung der Transformation zwischen Kamera und Laser wird in Kapitel 3.1 verdeutlicht.

2.2.4 Pose von Sensorträger und Kamera

Die Bestimmung der Sensorträgerpose¹ bezüglich des lokalen Weltkoordinatensystems erfolgt mittels der Flugzustandsschätzung unter Verwendung eines Navigationsfilters. Die Eingangsdaten dieses Filters liefern DGPS², IMU³ und Magnetometer⁴. Weitere Informationen zum verwendeten Navigationsfilter werden von Lorenz [60] präsentiert.

Die statische Pose der Kamera gegenüber dem Träger wird manuell nach der Montage bestimmt. Konstante Abweichungen, die durch Messungenauigkeiten und dynamische Abweichungen, welche durch Vibrationen während des Fluges entstehen, wurden als vernachlässigbar betrachtet.

Die Auswertung der Sensordaten im lokalen Weltkoordinatensystem setzt voraus, dass die Pose des Sensorträgers zum Zeitpunkt der Aufnahme der Sensordaten bekannt ist. Der Aufnahme- und Bereitstellungszeitpunkt der Posenschätzung durch den Navigationsfilter sind nicht identisch. Aus diesem Grund wird die Pose des Sensorträgers, die zum Zeitpunkt t_s der Aufnahme vorliegt, auf Basis der aktuellsten Lösung t_{nav} des Navigationsfilters approximiert. Die Zeitpunkte t_s und t_{nav} basieren auf einer synchronisierten Systemzeit, und es gilt $t_{nav} < t_s$. Zugrundeliegend auf den zum Zeitpunkt t_{nav} bestimmten Navigationsdaten, $\Phi_{ob,t_{nav}}$, $\dot{\Phi}_{ob,t_{nav}}$, $\mathbf{t}_{ob,t_{nav}}$, $\dot{\mathbf{t}}_{ob,t_{nav}}$ und $\ddot{\mathbf{t}}_{ob,t_{nav}}$ sowie der Zeitspanne $\Delta t = t_s - t_{nav}$ wird die approximierete Pose für den Zeitpunkt t_s wie folgt

$$\begin{aligned}\mathbf{t}_{ob,t_s} &= \mathbf{t}_{ob,t_{nav}} + \Delta t \dot{\mathbf{t}}_{ob,t_{nav}} + 0.5(\Delta t)^2 \ddot{\mathbf{t}}_{ob,t_{nav}}, \\ \Phi_{ob,t_s} &= \Phi_{ob,t_{nav}} + \Delta t \dot{\Phi}_{ob,t_{nav}},\end{aligned}\quad (2.13)$$

bestimmt.

2.3 Sensoren zur Umgebungserkennung

Für die dreidimensionale Wahrnehmung von Hindernissen entlang der Hauptbewegungsrichtung eines Fahrzeuges können eine Vielzahl von Sensoren eingesetzt werden. Eine Auswahl der zur Verfügung stehenden Sensoren sind beispielsweise Sonar-, Radar-, Laser- und Kamera-Systeme. Es

¹ Pose ist die Kombination von Position und Orientierung eines Objekts im dreidimensionalen Raum, DIN EN ISO 8373.

² Schätzung der absoluten Position und translatorischen Geschwindigkeiten auf Basis einer satellitengestützten Laufzeitenmessung in Kombination mit mindestens einer statischen Bodenstation. Der Verwendung der statischen Bodenstation ermöglicht eine Verbesserung der Genauigkeit gegenüber einer reinen Satellitenlösung.

³ Sensoren zur Messung von Beschleunigungen und Drehraten

⁴ Detektion des Erdmagnetfeldes zur Schätzung von Inklination und Deklination

erfolgt ein Überblick über vier Sensorarten. Darauf aufbauend wird auf die Eigenschaften von Laserscanner und monokularer Kamera, welche für das weitere Verständnis der Arbeit notwendig sind, tiefer gehend eingegangen.

2.3.1 Überblick

Sonar hat sich in der Nahbereichshindernisdetektion, innerhalb als auch außerhalb von Gebäuden durchgesetzt. Sonar- bzw. Ultraschallsensoren arbeiten mit akustischen Wellen und sind als aktiver und passiver Sensor einsetzbar und basieren auf der Ausbreitung von Schallwellen in einem Trägermedium.

Die Leistungsfähigkeit eines Sonars hängt stark vom Trägermedium ab in welchem es eingesetzt wird. Bei einem Einsatz in Luft sind Reichweiten von bis zu zehn Meter mit einer Auflösung im Zentimeter-Bereich möglich. Auf Grundlage der geringen Reichweite kommen Sonare in der unbemannten Luftfahrt häufig bei der Ermittlung der Höhe über Grund bei Landeverfahren in der letzten Phase der Landung zum Einsatz [21]. Die geringe Reichweite des Sensors schließt einen Einsatz bei der Detektion von Objekten in der Hauptflugrichtung allerdings aus.

Radar ist einer der wichtigsten Sensoren für die Kollisionsvermeidung in der bemannten Luft- und Raumfahrt und damit ebenso für den Einsatz in UA prädestiniert. Das Radar ist ein aktiver Sensor, welcher mit elektromagnetischen Wellen arbeitet [95] und traditionell eher für die Fernerkundung eingesetzt wird.

Ein Vorteil des Radars ist seine „all Wetter“-Tauglichkeit, welche es ermöglicht, besonders bei schlechten Sichtbedingungen hervorragende Ergebnisse bei der Bestimmung von Tiefeninformationen zu erzielen [26]. Die hervorragenden Leistungen bei widrigen Umweltbedingungen setzen eine vergleichsweise große Leistungsaufnahme und Leistungselektronik voraus. Dieses führt dazu, dass Radarsysteme im Vergleich zu anderen abbildenden Umweltsensoren vergleichsweise groß und schwer sind und viel Energie benötigen. Beim aktuellen Stand der Technik ist der Einsatz auf relativ große Luftfahrzeuge begrenzt. Der Trend zur Miniaturisierung könnte in Zukunft aber durch aus auch den Einsatz auf kleinen UAs ermöglichen.

Laserscanner bzw. engl. Light Detection and Ranging (LiDAR) gehört ebenfalls zu den aktiven Sensoren und arbeitet über die gezielte Ausleuchtung der Zielumgebung mittels Licht und der Detektion der auftretenden Reflexion auf möglichen Objektoberflächen.

LiDAR-Systeme sind von ihrer Leistungsfähigkeit beim aktuellen Stand der Technik zwischen Sonar und Radar einzuordnen. Die Reichweite von LiDAR liegt im Medium Luft im Allgemeinen über der von Sonar. Die Reichweite von Radar und dessen „all Wetter“-Tauglichkeit kann aber speziell bei einer hohen Partikeldichte im Trägermedium nicht erreicht werden. Die aktuell erreichten Reichweiten, die unabhängig von externer Beleuchtung sind sowie die kompakte Bauform machen LiDAR sehr interessant für viele Ansätze mit UA [36, 46, 55, 65, 80, 96].

Kameras passiver Bauart, die zur dreidimensionalen Hindernisdetektion eingesetzt werden, können sowohl monokular als auch multiokular aufgebaut sein. Eine weit verbreitete Variation eines multiokularen Ansatzes ist der Stereokameraaufbau [7, 45]. Stereoskopie erlaubt es mittels Triangulation eine indirekte Entfernungsmessung zu bestimmen. Indirekte Entfernungsmessungen können ebenfalls mit einer monokularen Kamera ermittelt werden, wenn die extrinsische Transformation zwischen den Aufnahmen bekannt und die betrachtete Szene statisch ist. Die indirekte als auch die multiokulare Entfernungsmessung enthält systematische Unzulänglichkeiten, die zu einer quadratisch abnehmenden Entfernungsauflösung mit zunehmender Entfernung führen. Ein weiteres Fehlerpotenzial enthalten inkorrekte Korrespondenzzuweisungen, die zu falschen Entfernungsmessungen führen. Basierend auf den Mängeln, wird den Entfernungsmessungen die mittels Stereoskopie entstanden sind, ein geringeres Vertrauen als bspw. Lasermessungen zugesprochen [46]. Aktive Kamerasysteme wie PMD⁵, die etwa eine Laufzeiten-Messung verwenden, sind beim aktuellen Stand der Technik noch nicht für die Hindernisdetektion an UA geeignet. Der Arbeitsbereich derartiger Systeme ist stark beschränkt. Eine Erweiterung des Arbeitsbereiches kann zu Mehrdeutigkeiten führen, die mittels alleiniger Nutzung der PMD-Kamera nicht aufgelöst werden können.

2.3.2 LiDAR

Im folgenden Abschnitt erfolgt die grundlegende Beschreibung der Sensoreigenschaften eines Laserscanners (LiDAR, LRF) zur Entfernungsmessung. Die Beschreibung beschränkt sich auf eine Entfernungsmessung mittels gepulster Laufzeitenmessung.

⁵engl. Photonic Mixer Device

Laserstrahl

Ein von einem Laser emittierter Lichtstrahl entspricht nicht einem mathematischen Strahl. Der Laserstrahl ist ein Strahlenbündel, das nach dem Verlassen des Lasers divergiert, siehe Abbildung 2.3. Der Durchmesser des

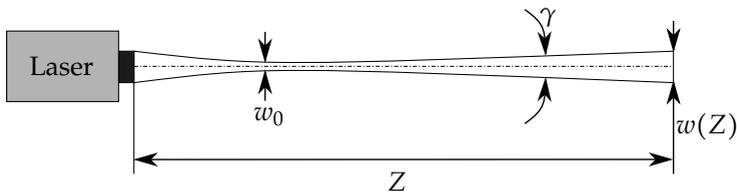


Abbildung 2.3: Strahlendivergenz bei Laserstrahlen

Strahls nimmt mit zunehmender Entfernung Z von der Strahlenquelle zu, wobei der Zusammenhang hyperbolisch ist. Der minimale Durchmesser des Strahls wird als Strahlentaille w_0 bezeichnet. Eine kleinere Strahlentaille führt zu einem größeren Divergenzwinkel γ . Der Strahlendurchmesser für einen Strahl der Wellenlänge λ kann für Z mit

$$w(Z) = w_0 \sqrt{1 + \left(\frac{\lambda Z}{\pi w_0^2} \right)^2} \quad (2.14)$$

bestimmt werden.

Im Fernfeld, für große Z , ist die Divergenz des Strahlenbündels fast proportional zu Z und die Äste der Hyperbel nähern sich zwei Geraden an. Die Geraden schließen den Winkel γ ein. Für den Durchmesser $w(Z)$ des Laserspots auf einem angestrahlten Objekt im Fernfeld gilt

$$w(Z) = 2Z \tan\left(\frac{\gamma}{2}\right). \quad (2.15)$$

Laufzeitenmessung

Für die Laufzeitenmessung mittels Laser werden eine Lichtquelle bzw. ein Transmitter, eine Senke bzw. ein Detektor sowie eine Elektronikeinheit, die eine Zeitdifferenz bestimmt, verwendet. Ein Aufbau wird in Abbildung 2.4 skizziert. Die Funktionsweise einer gepulsten Entfernungsmessung stellt sich wie folgt dar. Der Transmitter emittiert einen Laserpuls in Richtung des Objekts und teilt den Zeitpunkt der Transmission der Zeitmesseinheit

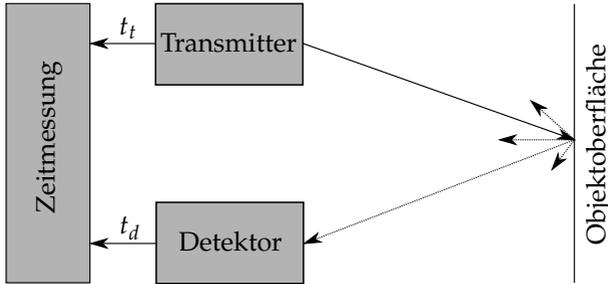


Abbildung 2.4: Prinzipieller Aufbau für eine Laufzeitenmessung

mit. Trifft der ausgesandte Strahl ein Objekt, wird er diffus reflektiert. Eine diffuse Reflexion findet statt, wenn die Oberfläche des Objektes eine große Rauheit relativ zu Wellenlänge λ aufweist. Die Rückstrahlung, auch als Echo bezeichnet, wird vom Detektor aufgefangen und der Zeitpunkt der Wahrnehmung wird wiederum dem Zeitmessmodul mitgeteilt. Dem Zeitmessmodul sind sowohl der Zeitpunkt der Transmission t_t als auch der Detektion t_d bekannt. Mittels der Zeitspanne

$$\Delta t = t_d - t_t \quad (2.16)$$

lässt sich unter Verwendung der Lichtgeschwindigkeit c die Entfernung

$$Z = \frac{\Delta t \cdot c}{2} \quad (2.17)$$

zwischen Objekt und Laser herleiten.

Wie aus 2.17 ersichtlich, hängt die Messgenauigkeit der Entfernungsmessung allein von der Genauigkeit der Zeitmessung ab. Die Messgenauigkeit ist unabhängig von der Entfernung, welche zwischen LiDAR und Objektoberfläche vorliegt.

Aus Gründen der Vollständigkeit sind noch die Verfahren Frequenzdifferenz oder Amplitudendifferenz zu nennen, mit denen ebenfalls die Entfernung zwischen Laser und Objekt bestimmt werden kann. Auf diese Verfahren soll aber nicht weiter eingegangen werden. Information dazu finden sich u.a. bei Adam [1].

Intensitätsmessung

Einige Laserscanner [37, 83, 93] messen neben Entfernungen je nach Modell auch Intensitäten des eingehenden Echos. Die Intensität des Echos, das ausgehend von einem emittierten Strahl, dessen Remission an einer Objektoberfläche wieder den Laserscanner erreicht, ist unter anderem von der Entfernung Z zwischen Objektoberfläche und Laser, dem Ausmaß der Fernfelddivergenz, der Oberflächenstruktur und dem Eintrittswinkel auf der Oberfläche abhängig.

Unter der Annahme, dass der verwendete Laserstrahl einer Strahlaufweitung gemäß einer Fernfelddivergenz (Gleichung: 2.15) unterliegt, erfolgt eine Verteilung der Strahlenergie auf einem Kugelsegment, das proportional mit dem Quadrat der Entfernung Z zwischen Objekt und Laserquelle wächst. Die Intensität I^6 nimmt folglich mit $I \propto 1/Z^2$ ab.

Die Oberflächenstruktur bzw. -gestalt des angestrahlten Objekts wirkt über den Grad der Remission auf die empfangende Intensität ein. Die Remission beschreibt eine diffuse Reflexion von Wellen im speziellen Licht. Der Remissionsgrad ρ gibt Auskunft über das Verhältnis von remittierter zu eingegangener Energie auf der Objektoberfläche, wobei $0 \leq \rho \leq 1$. Beispiele für die Remission einiger Materialien zeigt die Tabelle 2.1. Ein Wert von 10% entspricht $\rho = 0.1$. Für den Einfallswinkels Θ , welcher sich zwischen

Material/ Farbe	Remission(%)
Karton, matt schwarz	≈ 10
Karton, grau	≈ 20
Baumstamm	≈ 40
Papier, matt weiß	≈ 80
Aluminium oder Stahl, nicht oxidiert	$\approx 110 - 150$
Reflektor	> 2000

Tabelle 2.1: Beispiele für die Remissionsgrade verschiedener Materialien und Farben [82, S. 8].

Laserstrahl und der Normalen der Objektoberfläche aufspannt, wird davon ausgegangen, dass mit zunehmendem Θ weniger Licht diffus in die Richtung der Quelle remittiert wird [39]. Der Winkel Θ ist limitiert durch

⁶Leistung pro Fläche

$0^\circ \leq \Theta \leq 90^\circ$. Zusammengefasst werden die drei Einflussgrößen der Intensität I in folgender Gleichung

$$I \propto \frac{\rho \cos \Theta}{Z^2}. \quad (2.18)$$

Neben den drei präsentierten Einflussfaktoren hat auch die Art und Weise der Echo-Auswertung, die von jedem Laser durchgeführt wird, einen Einfluss auf die vom Laser bereit gestellte Intensität.

Auslenkung des Lasers

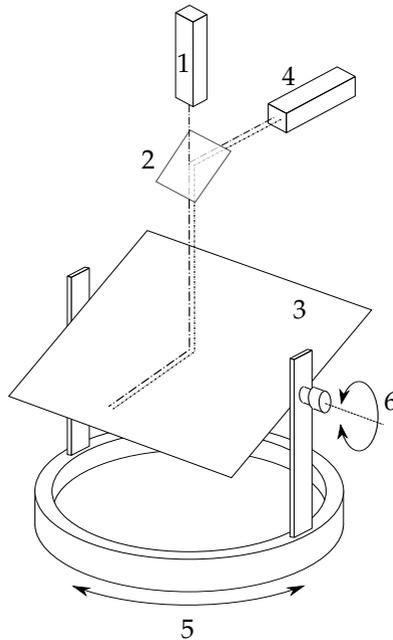
Eine einzelne Messung ist hinreichend, um den Abstand zu einem gewünschten Punkt zu bestimmen. Um die Form und Größe eines Objektes zu erhalten, sind im Allgemeinen mehrere Messungen notwendig. Laserscanner können mehrere Messungen an unterschiedlichen Positionen im Raum durch die Auslenkung des emittierten Strahls erreichen. Zur Auslenkung des Strahls wird entweder mittels einem oder zweier Spiegel der Strahl in die gewünschte Richtung abgelenkt, oder es wird über Servomotoren direkt die Ausrichtung der Strahlenquelle manipuliert.

Die Entfernungsmessung bzw. die Laufzeit des Lichts ergibt den ersten Freiheitsgrad. Für eine unkorrelierte dreidimensionale Abtastung werden zwei weitere Bewegungsachsen benötigt. Abbildung 2.5 zeigt exemplarisch einen derartigen Aufbau eines dreidimensionalen LiDAR.

Basierend auf diesem Aufbau in Abbildung 2.5 werden die kartesischen Messungen innerhalb des LiDAR-Koordinatensystem l mittels Kugelkoordinaten

$$\begin{aligned} x &= r \sin \Theta \cos \Phi \\ y &= r \sin \Theta \sin \Phi \\ z &= r \cos \Theta \end{aligned} \quad (2.19)$$

hergeleitet. Der Radius r innerhalb der Kugelkoordinaten entspricht der Entfernung Z aus dem vorhergehenden Abschnitt 2.3.2. Abbildung 2.5 zeigt ein Beispiel für eine Konstruktion, die ein dreidimensionales Abtasten der Umgebung erlaubt. In der Praxis umgesetzte Systeme beschränken sich oft auf ein zweidimensionales Scanfeld. Diese Systeme werden u.a. als Einzelenscanner oder monogoner Scanner [75] bezeichnet. Monogone Laserscanner nutzen einen flachen Spiegel zur Ablenkung der Strahlen, wobei die Normale der Spiegeloberfläche nicht parallel zum Strahlenverlauf zwischen Laserquelle und Spiegel ist. Des Weiteren rotiert der Spiegel, wobei die Rotationsachse parallel zu dem Strahlenverlauf zwischen Laserquelle



- | | |
|-------------------|--------------------------------------|
| 1. Transmitter | 5. vertikale Rotationsachse Θ |
| 2. Strahlenteiler | 6. horizontale Rotationsachse Φ |
| 3. Spiegel | ---- ausgesendeter Strahl |
| 4. Detektor | reflektierter Strahl |

Abbildung 2.5: Ablenkung eines Laserstrahls mit zwei rotatorisch gelagerten Spiegeln.

und Spiegel verläuft. Durch diesen Aufbau entsteht ein zweidimensionales Sichtfeld. Ein Messpunkt \mathbf{p}_l setzt sich aus dem Rotationswinkel Φ des Spiegels und der Entfernung z zusammen. Alle Strahlen und somit auch die Messungen liegen innerhalb einer Ebene $\mathbf{E} : \mathbf{p}_l(z, \Phi)$. Bezogen auf die Abbildung 2.5 ist $\Theta = \text{const.}$

Punktwolke

Mit den präsentierten Scannern, ob zwei- oder dreidimensional, können verschiedenste Szenen abgescannt werden, was je nach Dichte der Abtastung und Größe des Scanfeldes eine Vielzahl von Messungen ergibt. Eine zusammenhängende Ansammlung der Messungen wird als Punktwolke

$\mathbf{C} : p(x, y, z) \in \mathfrak{R}$ bezeichnet. Das einmalige Abfahren des gesamten Sichtfeldes wird oft als Scan bezeichnet. Ein Scan A enthält eine Punktwolke \mathbf{C} , welcher im Allgemeinen einem Erstellungsdatum t zugeordnet ist.

Ein Laserscanner tastet das zur Verfügung stehende Scanfeld seriell ab. Das heißt, dass jede Orientierung (Φ, Θ) des Strahls separat angefahren wird und die einzelnen Messungen nacheinander erfolgen, bis das komplette Scanfeld mit der definierten Winkelauflösung abgetastet ist. Innerhalb der Zeitspanne Δt , die der Scanner benötigt, um von einer Orientierung in die nächste zu verfahren, können von außen Bewegungen auf den Scanner aufgebracht werden. Die Bewegungen können bspw. durch einen Roboter, ein Fahrzeug oder ein Flugzeug erzeugt werden, an dem der Scanner montiert ist. Der Sensorträger bewegt sich gegenüber dem Weltkoordinatensystem, was ebenfalls zu einer Posenänderung des Scanners bezüglich des lokalen Weltkoordinatensystems führt. Die von dem Scanner abzutastenden Objekte befinden sich ebenfalls im Weltkoordinatensystem. Die Kombination der einzelnen Lasermessungen direkt im Koordinatensystem des Lasers mittels einer Punktwolke \mathbf{C}_t sowie eine Transformation von \mathbf{C}_t zum Zeitpunkt t in das lokale Weltkoordinatensystem unter Kenntnis der Sensorträgerbewegung berücksichtigt nicht die Bewegung, welche zwischen den einzelnen Messungen auftritt. Dieses führt zu einer Verfälschung von \mathbf{C}_t im Weltkoordinatensystem. Eine von der Bewegung bereinigte Punktwolke kann mittels einer Transformation

$$\left({}^0\mathbf{T}_{l,t_p}\right)^{-1} = \left({}^0T_{b,t_p}\right)^{-1} \left({}^bT_{l,t_p}\right)^{-1}, \quad (2.20)$$

jeder einzelnen Messung \mathbf{p}_l zum Zeitpunkt ihre Erstellung t_p in das von der Bewegung unabhängige lokale Weltkoordinatensystem erreicht werden. Die Transformation erfolgt mittels der Parameter Φ_{ob,t_p} und \mathbf{t}_{ob,t_p} des Laserscanners, sowie der Approximationen \mathbf{t}_{ob,t_p} und Φ_{ob,t_p} des Navigationsfilters 2.13, die zum Zeitpunkt \mathbf{p}_l vorliegen.

Abbildung 2.6 zeigt diesen Zusammenhang beispielhaft anhand eines Laserscanners, der sich translatorisch entlang seiner z-Achse auf ein Objekt O_o zu bewegt. Das Objekt ist bezüglich des lokalen Weltkoordinatensystems definiert. Die Transformation ${}^0\mathbf{T}_{l,t_p}$ zwischen Laserscanner und lokalem Weltkoordinatensystem ist zu jedem Zeitpunkt t_p ideal bekannt. Während der Bewegung erfolgen n -Messungen über eine Zeitspanne von $(n-1)\Delta t$, wobei Δt die Zeitspanne zwischen zwei Messungen beschreibt. Wird jede Messung \mathbf{p}_l zum Zeitpunkt ihrer Erzeugung in das lokale Weltkoordinatensystem transformiert und dort zu einer Punktwolke \mathbf{C} aus den Messungen $\mathbf{p}_{o,t}$ bis $\mathbf{p}_{o,t+(n-1)\Delta t}$ zusammengesetzt, ergibt sich die durchgehend schwarz gezeichnete Objektfläche. Werden die Messungen dagegen direkt im Laserkoordinatensystem zu einer Punktwolke \mathbf{C} kombiniert

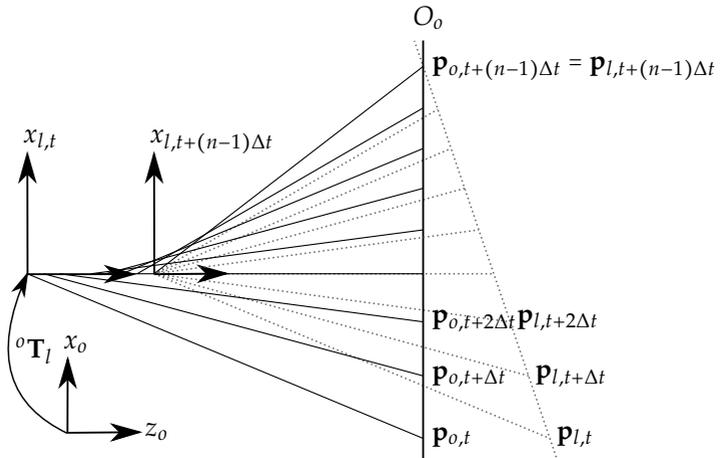


Abbildung 2.6: Bewegungsbereinigte Laserscannermessungen durch eine Transformation der Daten in das lokale Weltkoordinatensystem.

und zum Zeitpunkt $t + (n - 1)\Delta t$ mit den zu diesem Zeitpunkt vorliegenden Parametern des Navigationsfilters in das lokale Weltkoordinatensystem transformiert, ergibt sich die grau gepunktete Objektoberfläche. Die unterschiedliche Abbildung von O_0 ist somit auf die unterlassene Berücksichtigung der Bewegung innerhalb der Laserkoordinaten zurückzuführen.

Umweltbedingte Detektionsfehler

Der von dem Laserscanner ausgesendete Lichtstrahl entspricht, wie erläutert, nicht einem ausdehnungslosen Strahl im mathematischen Sinne. Der Querschnitt des Strahls bedeckt eine Fläche

$$A(Z) \approx \pi \cdot \left(\frac{w(Z)}{2} \right)^2, \quad (2.21)$$

deren Ausdehnung durch die Entfernung zwischen Quelle und beleuchtetem Objekt bedingt ist. Je nach Größe der Fläche und des Eintrittswinkels des Lichtstrahls auf einer Objektoberfläche variiert der zeitliche Abstand zwischen der ansteigenden und abfallenden Signalflanke des reflektierten Echos. Es ergibt sich pro Laserpuls aber ausschließlich ein Echo, das in seiner Länge variiert. Basierend auf der Ausdehnung eines Strahls kann

die Situation eintreten, dass mehrere Objekte durch einen einzelnen Strahl beleuchtet werden. Bei der simultanen Bestrahlung bspw. zweier Objekte unterschiedlicher Entfernungen ergeben sich zwei Echos, die basierend auf einem gemeinsamen Quellimpuls den Detektor des Laserscanners unabhängig, zu unterschiedlichen Zeitpunkten, erreichen. Die Laufzeitmessung ergibt für das Objekt im Vordergrund eine Zeitspanne Δt_f und für das Objekt im Hintergrund eine Zeitspanne Δt_b , wobei $\Delta t_f < \Delta t_b$ ist. Das beschriebene Phänomen wird von Böhler [15] als „Edge Effect“ bezeichnet.

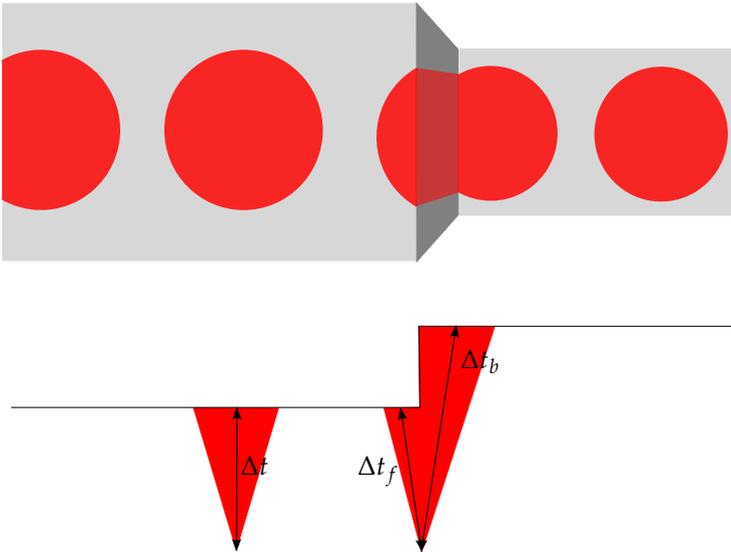


Abbildung 2.7: Edge Effect: ein ausgesendeter Laserpuls generiert durch die simultane Beleuchtung von zwei Objekten, zwei Echos.

In der Praxis [37, 83] sind mehrere Verfahren umgesetzt, wie Multiechos behandelt werden. Das einfachste Vorgehen ist eine ausschließliche Nutzung eines der Echos. Aus konservativen Sicherheitsgründen wird die Nutzung des ersten Echos bevorzugt, welches das nächstliegende Objekt zur Strahlenquelle beschreibt. Ein weiteres Vorgehen ist die Bildung des Mittelwerts über mehrere Echos. Dieses Vorgehen führt zu einer gemittelten Entfernungsmessung, die auch als „Mixed Pixel“ [40, 71] bezeichnet wird. Problematisch bei diesem Vorgehen ist, dass nicht das nächstliegende Objekt angezeigt wird und sich zwischen dem berechneten Strahlenende und der Laserquelle somit „unbetrachtete“ Objekte befinden können.

Die Detektion des Laserlichts basiert auf einer diffusen Reflexion an einer Oberfläche, deren Rauheit größer der Wellenlänge des reflektierten Lichts ist. Bei Oberflächen, wie Spiegel oder Wasser, die diese Anforderungen nicht notwendig erfüllen, treten gerichtete, totale Reflexion und Brechung auf. Das Licht wird nicht immer zur Quelle reflektiert und es erfolgt keine Signalisierung der Oberflächen.

Unterschiedlich farbige und reflektive Oberflächen führen bei identischer, konstanter Entfernung zu variierenden Entfernungsmessungen. Bei farbigen Oberflächen ist dies auf das Zusammenspiel zwischen Wellenlänge des Laserlichts und der Oberflächen zurückzuführen [71]. Unterschiedliche Graufärbungen bzw. Materialien haben unterschiedliche reflektive Eigenschaften, die die reflektierte Lichtmenge reduzieren können und damit ebenfalls Einfluss auf die Eigenschaften des Echos haben [15]. Von Herbert [40] wird die Beeinflussung von Entfernungsmessungen durch die Intensität als „Range/Intensity crosstalk“ bezeichnet. Der Einfluss von Farbe und Reflektivität auf die Hindernisdetektion ist allerdings vernachlässigbar.

Die Reflexion von Licht an Oberflächen beinhaltet sowohl gerichtete als auch diffuse Bestandteile. Der Eintrittswinkel des Lichts auf die Objektfläche hat einen Einfluss auf die Menge der Energie, die an die Quelle zurückgestrahlt wird. Je kleiner der Eintrittswinkel ist, desto größer ist die Messgenauigkeit, welche mit zunehmendem Winkel abnimmt. Okubo [71] zeigt in einem Beispiel, dass ab einem Eintrittswinkel größer $\pm 30^\circ$ keine sinnvolle Detektion mehr möglich ist.

Das Auftreten von Abweichungen, welche auf die Objekteigenschaften zurückzuführen sind, kann im Vorfeld bspw. bei einer Kalibrierung nicht berücksichtigt werden.

Interpretation von Laserdaten

Die Daten, die ein Sensor bereitstellt, entsprechen nur bedingt der Realität. Externe Einflüsse und Fehler erzeugen Abweichungen vom Ideal. Für eine realistische Einschätzung der approximierten Messung ist eine Berücksichtigung der entsprechenden Einflussfaktoren notwendig. Die jeweiligen Einflussfaktoren sind vom Sensortyp abhängig und können in einem Sensormodell kombiniert werden.

Sensormodelle werden in direkte bzw. physikalische Sensormodelle und inverse Sensormodelle unterteilt [90]. Ein physikalisches Sensormodell beschreibt die erwartete Messung bzw. die Wahrscheinlichkeitsverteilung einer Messung auf Basis eines bekannten Objektes. Voraussetzung sind die spezifischen Eigenschaften wie bspw. Reichweite, Auflösungsvermögen, Rauschverhalten und Messmethodik. Mittels eines inversen Sensormodells

wird, ausgehend von einer gegebenen Sensormessung p , auf die Wahrscheinlichkeit eines Objektes an der prognostizierten Position geschlossen. Die Herleitung eines inversen Sensormodells basiert auf dem physikalischen Sensormodell des jeweiligen Sensors. Thrun et. al [88] nutzten zur Herleitung eines inversen Sensormodells eine Zufallskarte, von der mittels eines direkten Sensormodells Messungen abgeleitet werden. Die abgeleiteten Messungen bilden die Grundlage für Wahrscheinlichkeitsaussagen über die Gestalt der Ausgangskarte.

Dieses Vorgehen ist aufwendig, da auf einer großen Strichprobenbasis operiert wird, um repräsentative Ergebnisse zu erreichen. Ein pragmatischeres Vorgehen schlagen Elfes [24] und Elsner [25] vor. Aufbauend auf der Annahme der klassischen Physik, „Wo ein Körper ist kann kein weiterer sein“ bzw. einer ungehinderten, geradlinigen Sicht vom Sensorursprung auf den Objektpunkt p , wird von einem freien Raum zwischen beiden Punkten ausgegangen. Dieser Objektpunkt stellt eine Unterbrechung des Strahls dar, was auf die Belegung des Raumes mit einem Objekt schließen lässt.

Das Vertrauen, das den Entfernungsapproximationen des Fusionsensors entgegengebracht werden kann, beruht auf dem Vertrauen, das die Einzelsensoren bereitstellen.

Anhaltspunkte für das jeweilige Vertrauen, das in die Messergebnisse der einzelnen Sensoren gesetzt werden kann, finden sich bei den jeweiligen Messmethoden und deren physikalischen Eigenschaften.

Physikalisches Sensormodell Der Laserscanner generiert Entfernungsmessungen, die dreidimensional verortet werden können. Ein Sensormodell, das verdeutlicht, mit welcher Wahrscheinlichkeit eine tatsächliche Entfernungsmessung von der erwarteten Messung abweicht, wird von Fox [28], Burgard und Thrun [87] vorgestellt. Ursachen für die Differenzen zwischen der Erwartung und der tatsächlicher Messung sind Messauflösung, -rauschen, unerwartete Objekte, übersehene Objekte und zufällige Messungen.

Die Abweichungen, die auf Messauflösung und -rauschen zurückzuführen sind, werden von Fox et al. durch eine Gauß-Verteilung umgesetzt, wobei der Erwartungswert der erwarteten Objektentfernung entspricht. Die Auflösung bzw. das Rauschen kann unter anderem durch externe Einflussfaktoren wie Beleuchtung, Oberflächenstruktur, Material oder Entfernung variieren.

Fox et al. verwenden die Entfernungsmessungen für die Kartierung der Umwelt. Bei der Kartierung wird sich vornehmlich auf statische Objekte konzentriert. Die Aufzeichnungen der Messungen erfolgen allerdings in einer dynamischen Umwelt, was dazu führt, dass unerwartete, dynami-

sche Objekte die Messungen kreuzen. Dynamisches Objekt können Personen, Fahrzeuge aber auch kleine Objekte, wie Regentropfen oder Staub, sein. Die Detektion von dynamischen Objekten führt zu Abweichungen gegenüber der erwarteten statischen Objektkulisse.

Neben der Detektion von unerwarteten Objekten können erwartete Objekte aufgrund ihrer Oberflächeneigenschaften übersehen werden. Ausgehend von einem Laser können Objektoberflächen mit optischen Eigenschaften Einfluss auf den Strahlenverlauf des Lasers nehmen. Eine Absorption oder eine gerichtete Reflexion eines Strahls verhindern eine diffuse Reflexion oder führen zu einer Reflexion an nachgelagerten Oberflächen. Oberflächen mit derartigen Eigenschaften sind unter anderem Glas, Wasser bzw. Oberflächen, deren Rauheit unter der Wellenlänge des eingesetzten Lichts liegt.

Führen die Oberflächeneigenschaften zu einer gerichteten Reflexion, kann ein abgelenkter Strahl ein anderes Objekt treffen. Die Folge kann eine diffuse Reflexion an der Oberfläche des zweiten Objekts sein, dessen Ergebnis ein Echo ist, das den Laserscanner erreicht. Dieser Vorgang wird als Mehrwegeausbreitung bezeichnet. Wird nach einer gerichteten Reflexion kein weiteres Objekt vor Erreichen der maximalen Messdistanz des Lasers getroffen, wird keine Messung aufgezeichnet. Nicht detektierbar sind ebenso Oberflächen bei denen eine totale Reflexion des Lichts auftritt.

Fox et al. drücken die Wahrscheinlichkeit, dass eine initiierte Entfernungsmessung zu einem bekannten Objekt kein Echo generiert bzw. die bestimmten Messresultate außerhalb des maximalen Messbereiches des Sensors liegen, mit einer speziellen Gleichverteilung am äußersten Rand des Wahrnehmungsbereiches aus.

Reale, unerwartete Objekte beeinflussen einen Laserstrahl ausschließlich, wenn sie sich zwischen der Laserquelle und dem erwarteten Objekt befinden. Fox et al. berücksichtigen unerwartete Objekte ebenfalls hinter dem erwarteten Objekt. Die Detektion eines unerwarteten Objekts hinter dem erwarteten Objekt ist nur möglich, wenn der Strahl, der die Entfernung zwischen Laserquelle und dem naheliegenden Objekt angibt, ein Objekt durchdringt oder an seiner Oberfläche reflektiert wird. Die erste Variante tritt bei transparenten Objekten wie Glas oder Eis ein. Eine Reflexion hätte zur Folge, dass der Strahl an einem anderen Objekt reflektiert wird und die bestimmte Entfernung damit größer als die erwartete ist.

Fox et al. setzten zur Berücksichtigung von dynamischen Objekten und Mehrwegeausbreitung eine geometrische Verteilung an, die mit zunehmender Entfernung abnimmt und bei der maximalen Sensorreichweite endet. Schlussendlich berücksichtigen Fox et al. noch zufällige Messungen, für die es keine Erklärung gibt. Diese Messungen können bei jeder Ent-

fernung zur Laserquelle auftreten und sind gleichmäßig über die Sensorreichweite verteilt.

Abbildung 2.8 zeigt die Kombination der drei Verteilungen zu einem Sensormodell. Die Gauß-Verteilung in blau repräsentiert Sensorauflösung und Rauschen. Der rote Bereich zeigt eine geometrische Verteilung, welche für die Wahrscheinlichkeit steht, ein unbekanntes Objekt zu detektieren. Der grüne Abschnitt verdeutlicht die Wahrscheinlichkeit von unklaren Messungen mittels einer Gleichverteilung. Der Sprung in Höhe der maximalen Sensorreichweite steht für die Wahrscheinlichkeit, eine Messung zu generieren, deren Resultat außerhalb der Sensorreichweite liegt.

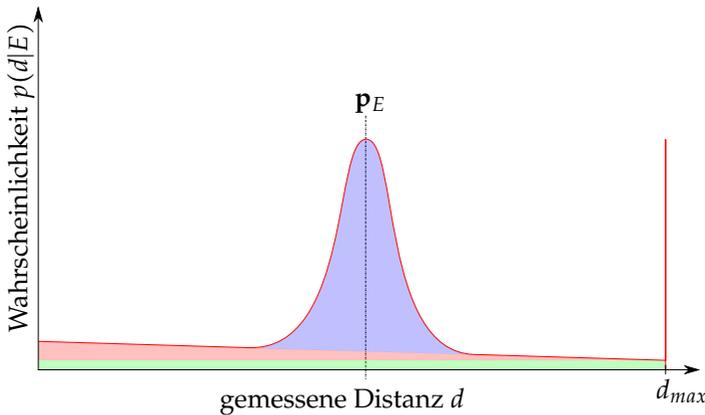


Abbildung 2.8: Kombiniertes Sensormodell nach Fox et al. [28].

Inverses Sensormodell Ein inverses Sensormodell wird eingesetzt, um auf Basis einer Sensormessung \mathbf{p} auf die Umgebung zu schließen. Von besonderem Interesse ist hierbei, ob Teile des Raumes frei oder von Objekten belegt sind. Ausgehend von einem Sensor, der mittels eines Strahls eine Entfernungsmessung $Z_{\mathbf{p}}$ erzeugt, wird Auskunft über die Belegung des von dem Strahl durchschnittenen Raums gegeben. Vorgestellt wurde das inverse Sensormodell von Elfes [24], wobei vom Idealfall einer Entfernungsmessung ausgegangen wird. Ausgehend von einer ungehinderter, gradlinigen Sicht vom Sensorursprung auf den Objektpunkt \mathbf{p} kann der Raum zwischen beiden Punkten als frei betrachtet werden. Der Objektpunkt stellt eine Unterbrechung des Strahls dar, was die Belegung des Raumes verdeutlicht. Die Unterbrechung des Strahls durch \mathbf{p} verhindert eine weitere Betrachtung des Raumes. Aufbauend darauf ist folglich keine Einordnung in die Kategorien frei bzw. belegt nach dem Durchlaufen von

\mathbf{p} möglich. Zur Vervollständigung des Zustandsraumes wird die Kategorie „unbekannt“ für den Raum nach dem Objektpunkt bis zur maximalen Sensorreichweite Z_{max} verwendet [7]. Über den Verlauf des Strahls ergibt sich die Belegungswahrscheinlichkeit

$$p(Z, Z_p) = \begin{cases} 0 & \text{für } 0 \leq Z < Z_p, \\ 1 & \text{für } Z = Z_p, \\ 0.5 & \text{für } Z_p < Z \leq Z_{max}, \end{cases} \quad (2.22)$$

welcher in Abbildung 2.9 visualisiert ist.

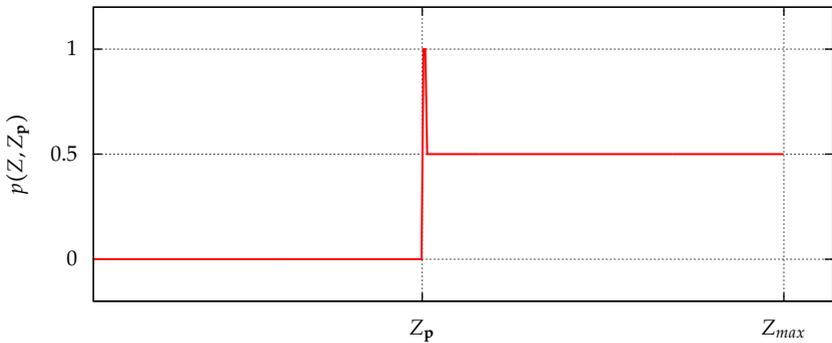


Abbildung 2.9: Inverses Sensormodell abgeleitet von einer idealen Sensormessung, nach Elfes [24].

Die Ausführungen in Abschnitt 2.3.2 verdeutlichen, dass die Anwendung eines idealen Sensormodells eine Vereinfachung darstellt. Für eine realistischere Berücksichtigung einer Messung wird das ideale Sensormodell um mögliche Fehlerquellen erweitert [46]. Aufbauend auf dem physikalischen Sensormodell wird die Signifikanz der Freiraumbewertung sowie der Messung selbst reduziert. Die Reduktion der Signifikanz spiegelt die Möglichkeit wider, dass Hindernisse „übersehen“ werden oder die erhaltende Messung auf einer Mehrwegeausbreitung beruht.

Die Belegtheitswahrscheinlichkeit $p(Z, Z_p)$ für ein realistisches, inverses Lasermodell ergibt sich wie folgt

$$p(Z, Z_p) = p_{base}(Z, Z_p) + p_{obj}(Z, Z_p). \quad (2.23)$$

Die Belegungswahrscheinlichkeit p_{base} ist die grundlegende Belegung der Zellen vor und nach der Detektion der Messung. Folglich ist

$$p_{base}(Z, Z_{\mathbf{p}}) = \begin{cases} p_{free} & \text{für } 0 \leq Z < Z_{\mathbf{p}}, \\ p_{unknown} & \text{für } Z_{\mathbf{p}} \leq Z < Z_{max}. \end{cases} \quad (2.24)$$

Für die Belegtheitswahrscheinlichkeit p_{free} werden Werte von $0 < p_{free} < p_{unknown}$ empfohlen, wobei $p_{unknown} = 0.5$ ist.

Die Berücksichtigung des Sensorrauschens bei der Belegungswahrscheinlichkeit der Messung erfolgt mittels einer Gauß-Verteilung durch

$$p_{obj}(Z, Z_{\mathbf{p}}) = \left(\eta_l \frac{1}{\sigma_l \sqrt{2\pi}} + p_{unknown} - p_{base}(Z, Z_{\mathbf{p}}) \right) e^{\left(-\frac{(Z-Z_{\mathbf{p}})^2}{2\sigma_l^2} \right)}. \quad (2.25)$$

Mittels des Parameters σ_l wird das Rauschen bzw. das Auflösungsvermögen des Sensors berücksichtigt. Für den Laser ist σ_l von der Entfernung unabhängig konstant. Der Parameter η_l wird zur Signifikanzreduktion der Messung \mathbf{p} verwendet. Für die Definition von η_l muss gelten,

$$\forall Z \geq Z_{min} : \{p(Z, Z_{\mathbf{p}}) \in \mathfrak{R} \mid 0 \leq p(Z, Z_{\mathbf{p}}) \leq 1\}. \quad (2.26)$$

Visualisiert ist Berechnung 2.23 in Abbildung 2.10.

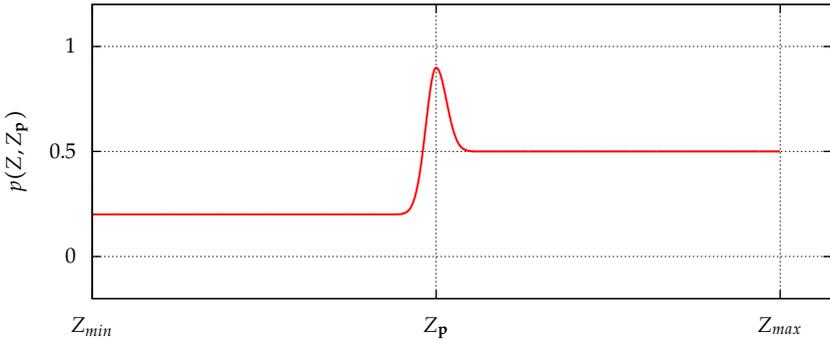


Abbildung 2.10: Inverses Sensormodell unter Berücksichtigung der von Fox et al. [28] dargelegten Fehlerquellen.

Tritt der Zustand $Z_{max} < Z_{\mathbf{p}}$ ein, gilt für den betrachteten Strahl

$$p(Z, Z_{\mathbf{p}}) = p_{free} \text{ für } Z < Z_{max}. \quad (2.27)$$

Für Entfernungen $Z_{max} \leq Z \leq Z_{\mathbf{p}}$ wird kein $p(Z, Z_{\mathbf{p}})$ ermittelt.

2.3.3 Monokulare Kamera

Lochkameramodell

Als Grundlage eines realen Kameramodells dient das Lochkameramodell. Das Lochkameramodell generiert eine abbildungsfehlerfreie, perspektivische, zentralisierte Abbildung des Blickfeldes auf die Bildebene mittels einer unendlich kleinen Blendenöffnung. Der Aufbau ist ein zentriertes System, wobei die Mittelpunkte der zentrierten, sphärischen Flächen auf einer Geraden, der optischen Achse, liegen. Alle Lichtstrahlen, die von Objekten auf die Bildebene verlaufen, schneiden sich im Projektionszentrum O_c auf der optischen Achse. Der Abstand zwischen O_c und dem Bildhauptpunkt H ist die Kamerakonstante c . Die Projektion $\mathbf{P} \in \mathbb{R}^{3 \times 4}$ eines dreidimensio-

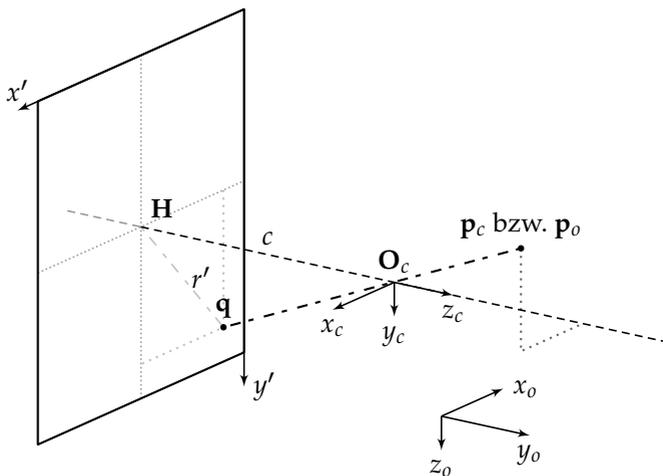


Abbildung 2.11: Lochkameramodell

nalen Punktes $\mathbf{p} = (x, y, z)$ aus dem System o in Bildebenenkoordinaten $\mathbf{q} = (x', y')$, beschrieben mit homogenen Koordinaten, ist für ein Lochkameramodell

$$\mathbf{q} = \mathbf{Pp}. \quad (2.28)$$

Die Projektion \mathbf{P} ist eine Kombination der intrinsischen und extrinsischen Orientierungsparameter. Die intrinsische Orientierung

$$\mathbf{K} = \begin{bmatrix} \frac{c}{s_x} & -\frac{c}{s_x} \cot(\Theta) & x_0 \\ 0 & \frac{c/s_y}{\sin(\Theta)} & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.29)$$

ist durch den inneren Aufbau der Kamera definiert und setzt sich aus der Position des Bildhauptpunktes $\mathbf{H}(x_0, y_0)$, der Kamerakonstanten c , der Scherung Θ der Bildachsen sowie der Abmessungen eines Bildpunktes in beiden Dimensionen (s_x, s_y) zusammen. Die extrinsischen Parameter setzen sich aus der Position \mathbf{t}_{oc} und Orientierung \mathbf{R}_{oc} des Kamerakoordinatensystems c im Weltkoordinatensystem o zusammen

$$\mathbf{P} = \mathbf{K} \mathbf{R}_{oc}^{-1} [\mathbf{I} | -\mathbf{t}_{oc}] \quad (2.30)$$

[61, S. 52]. Eine Abbildung eines Objektpunktes auf die Bildebene ohne Berücksichtigung der Rotation $\mathbf{R} = \mathbf{I}$ und Translation $\mathbf{t} = (0, 0, 0)^T$ der Kamera im Weltkoordinatensystem ergibt somit

$$\mathbf{q} = \mathbf{K} \mathbf{p}. \quad (2.31)$$

Bei der Abbildung auf die Bildebene geht die Information der Entfernung zwischen Bildebene und Objektpunkt dem Kamerakoordinatensystem verloren. Auf der Bildebene sind ausschließlich relative Abmessungen von Objekten bekannt. Aus diesem Grund ist eine Rückprojektion von der Bildebene auf den Objektpunkt allein mittels der Informationen, die ein Kamerabild bereitstellt, nicht möglich. Es können ausschließlich Aussagen zum Strahl gemacht werden, der ausgehend vom Objektpunkt den Bildpunkt generiert hatte. Wo sich der Objektpunkt auf dem Strahl befindet ist dagegen unbekannt. Monokulare Kamerabilder werden auch als skalierungsinvariant bezeichnet.

Aberrationen

Das Lochkameramodell ist ein symmetrisches optisches System und ermöglicht eine abbildungsfehler- bzw. aberrationsfreie Projektion. In der Praxis eingesetzte Kamerasysteme, die optische Systeme wie Linsen oder Spiegel verwenden, erfüllen diesen Idealfall nicht. Aus diesem Grund ist es notwendig, das Lochkameramodell um diese Komponenten zu erweitern.

Verzeichnung oder radial-symmetrische Aberrationen entstehen bei der Verwendung von asymmetrischen Optiken. Für derartige Optiken entspricht der Einfallswinkel τ in das optische System nicht dem Austrittswinkel τ' , was zu einer radialen Abweichung Δr auf der Bildebene führt.

Chromatische Aberrationen liegen vor, wenn ein frequenzabhängiger Brechungsindex des optischen Mediums zur Dispersion des Lichts führt. Dieser Vorgang generiert, abhängig von der Frequenz, unterschiedliche Brennpunkte F , was zu einer Reduktion der Bildschärfe führt. Monochromatisches Licht führt zu einer Abweichung, die der radial-symmetrischen Verzeichnung entspricht. Polychromatisches Licht führt zu einer radialen Verschiebung, welche je nach Frequenz unterschiedlich ausgeprägt ist (Regenbogen).

Sphärische Aberrationen führen zu einer Verschiebung Δf des Brennpunktes, abhängig von dem Einfallsradius. Der Einfallsradius ist der orthogonale Abstand zwischen dem einfallenden Licht und der optischen Achse im Objektraum. Mit zunehmendem Einfallsradius werden Strahlen stärker gebrochen, was zu einer Verschiebung von F' in Richtung des bildseitigen Projektionszentrums O' führt. Qualitativ hochwertige Objektive sind weitestgehend frei von sphärischer Aberration. Aus diesem Grund ist eine Berücksichtigung in der Modellierung nicht notwendig [61, S. 42].

Astigmatismus und Bildfeldwölbung sind zwei eng verknüpfte Aberrationen, welche für Strahlenbündel auftreten, deren Objektpunkt nicht auf der optischen Achse liegt und die schräg in die Linse einfallen. Der schräge Einfall des Bündels auf die brechende Fläche führt zu unterschiedlichen Winkeln bei der Brechung und generiert eine Abbildung, welche nicht in einen Bildpunkt sondern in einen elliptischen unscharfen Fleck resultiert. Eine Berücksichtigung dieser Aberrationen findet im Allgemeinen nicht statt, da davon ausgegangen wird, dass basierend auf entsprechenden Linsenkonstruktionen die entstehenden Abweichungen unterhalb der messbaren Größen liegen.

Koma ist eine Kombination von sphärischer Aberration und Astigmatismus, welche bei schräg zur optischen Achse einfallenden Strahlenbündeln auftritt und anstelle eines scharfen Bildpunktes einen Bildpunkt mit zum Bildrand gerichtetem „Schweif“ (griech. $\kappa\omicron\mu\alpha$ = Haar) generiert.

Bis zu diesem Punkt präsentierte Aberrationen basieren auf der Form der

Linse, wobei aber davon ausgegangen wird, dass es sich bei der Linse, um einen rotationssymmetrischen Körper handelt, dessen Rotationsachse auf der optischen Achse liegt. Dieses ist bei in der Praxis eingesetzten Optiken aber nicht immer der Fall. Schiefstellung und Dezentrierung der Linsen führen zu einer gekrümmten optischen Achse, wodurch radial-symmetrische Effekte von tangentialer und asymmetrischer Verzeichnung überlagert werden.

Neben der Lage und Position der Linsen sind eine Schiefstellung der Bildebene oder deren Achsen zueinander möglich. Eine nicht zur optischen Achse orthogonale Bildebene führt zu projektiven Abbildungsfehlern. Nicht orthogonale Bildkoordinatenachsen (Scherung) bzw. eine ungleichmäßige Skalierung dieser führen zu einer affinen Verzeichnung.

Die Abweichungen, die durch die Abbildungsfehler gegenüber dem idealen Lochkammermodell entstehen, werden mittels des Korrekturterms $\Delta\mathbf{q}(\Delta x, \Delta y)$ im Bildkoordinatensystem ausgedrückt. Die Verzeichnungen beziehen sich auf den Bildhauptpunkt \mathbf{H} , weswegen eine Normierung eines gemessenen, verzeichneten Bildpunktes $\mathbf{q}(x, y)$ gegenüber diesem durchgeführt wird. Der um den Bildhauptpunkt korrigierte Bildpunkt

$$\mathbf{q}^\circ(x^\circ, y^\circ) = \mathbf{q} - \mathbf{H} \quad (2.32)$$

und der dazugehörige Bildradius

$$r^\circ = \sqrt{x^{\circ 2} + y^{\circ 2}} \quad (2.33)$$

ermöglichen eine Berechnung der um die Verzeichnung korrigierten Bildkoordinaten

$$\mathbf{q}' = \mathbf{q}^\circ - \Delta\mathbf{q} \text{ bzw. } \mathbf{q}^\circ = \mathbf{q}' + \Delta\mathbf{q}. \quad (2.34)$$

Der Korrekturterm $\Delta\mathbf{p}'$ enthält radial-symmetrische $\Delta\mathbf{p}_{sym}$, asymmetrische $\Delta\mathbf{p}_{asym}$ und affine $\Delta\mathbf{p}_{aff}$ Elemente:

$$\Delta\mathbf{q} = \Delta\mathbf{q}_{sym} + \Delta\mathbf{q}_{asym} + \Delta\mathbf{q}_{aff}. \quad (2.35)$$

Der Anteil der radial-symmetrischen Verzeichnung lässt sich mittels

$$\begin{aligned} \Delta r^\circ_{sym} &= A_1 r^{\circ 3} + A_2 r^{\circ 5} + A_3 r^{\circ 7} - r^\circ (A_1 r_0^2 + A_2 r_0^4 + A_3 r_0^6) \\ \Delta\mathbf{q}_{sym} &= \mathbf{p}' \frac{\Delta r^\circ_{sym}}{r^\circ} \end{aligned} \quad (2.36)$$

bestimmen [19]. Die Verzeichnungsparameter A_1 bis A_n werden mittels Reihenentwicklung der Seidel-Reihe hergeleitet. Die tangentiale und radial-asymmetrische Verzeichnung wird durch

$$\begin{aligned} \Delta x_{asym} &= B_1 (r^{\circ 2} + 2x^\circ y^\circ) + B_2 x^\circ y^\circ \\ \Delta y_{asym} &= B_2 (r^{\circ 2} + 2y^\circ y^\circ) + B_1 x^\circ y^\circ \end{aligned} \quad (2.37)$$

beschrieben [19]. Der Bestandteil der Affinität und Scherung geht mittels

$$\begin{aligned}\Delta x_{aff} &= C_1 x^\circ + C_2 y^\circ \\ \Delta y_{aff} &= 0\end{aligned}\tag{2.38}$$

in die Gesamtverzeichnung ein [22].

Wie aus den Berechnungen ersichtlich ist, werden die verzeichneten und normierten Koordinaten \mathbf{q}' mittels der nicht verzeichneten Koordinaten \mathbf{q}'' bestimmt. Dieses Vorgehen ist nicht invertierbar, so dass eine Bestimmung von \mathbf{q}'' aus einem gemessenen Bildpunkt \mathbf{q}' nur iterativ möglich ist. Ein von Tsai und Lenz [57, 92] präsentierte Verfahren ermöglicht eine Invertierung, berücksichtigt allerdings nur die radial-symmetrische Verzeichnung. Dieses führt zu Defiziten bei der Genauigkeit [29].

2.4 Segmentierung

2.4.1 Überblick

Die Aufteilung eines Bildes bzw. die Gruppierung von Bildelementen wie Pixel oder Voxel⁷ zu zusammenhängenden Regionen wird in der Bildverarbeitung und Computergrafik als Segmentierung bezeichnet. Die Segmentierung erfolgt mittels eines Homogenitätskriteriums, welches die Elemente (Pixel, Voxel) in Mengen anforderungsbedingt gruppiert. Die Attribute einer Segmentierung sind:

- vollständig, jedes Pixel wird mindestens einem Segment zugeordnet,
- überschneidungsfrei, kein Pixel ist mehreren Segmenten zugeordnet und
- zusammenhängend, jedes Segment bildet eine zusammenhängende Region.

Formal ergibt sich für eine vollständige, überschneidungsfreie und zusammenhängende Segmentierung einer gegebenen Menge B mit den Elementen p_1, \dots, p_n und die gesuchten Segmente S_i folgende Darstellung:

$$\begin{aligned}\bigcup_{i=1, \dots, s} S_i &= B \\ S_i \cap S_j &= \emptyset \\ p_s \in S_j, p_t \in S_i : p_s \sim p_t &\Leftrightarrow i = j.\end{aligned}\tag{2.39}$$

⁷Gitterpunkt in einem dreidimensionalen Gitter. Vergleichbar einem Pixel in einem zweidimensionalen Gitter.

Die Segmentierung von B erfolgt mit dem Ziel, Objekte voneinander bzw. vom Hintergrund zu separieren, um einen objektbasierten Zugriff auf die einzelnen Segmente zu ermöglichen.

Der Raum, der zwei- bzw. dreidimensional abgebildet wird, kann eine große Anzahl von Objekten enthalten, die sich gegenseitig überlappen. Die Frage, die sich daraus ergibt, ist, welche Regionen gehören zu welchem Objekt bzw. wie homogen sind die Objektflächen bezüglich eines bestimmten Kriteriums.

Ansätze der Segmentierung können in vier grundlegende Konzepte unterschieden werden: pixel-, region-, kanten- sowie modellbasierte Verfahren [49, S. 449]. Die Entscheidung, welcher Ansatz zur Segmentierung gewählt wird, ist durch die Ausgangsmenge B und das gewünschte Ergebnis bedingt.

Pixelbasierte Verfahren sind einfach, flexibel und schnell. Für jedes Pixel wird eine von dessen Nachbarn unabhängige Einzelentscheidung gefällt, zu welcher Region das Pixel gehört. Dies führt nicht selten zu kohärenten Regionen, darüber hinaus ist nicht geklärt, in welchem Zusammenhang das einzelne Pixel zu der Gruppe steht. Pixel-basierte Ansätze weisen eine starke Anfälligkeit für Helligkeitsänderungen auf, was zu Schwächen bei dynamischen oder teilweisen Beleuchtungsänderungen führt. Das Ergebnis weist zusätzlich eine sehr starke Abhängigkeit von den Schwellwertparametern auf, was in dynamischen Umgebungen eine kontinuierliche Adaption nötig macht.

Regionbasierte Verfahren, auch als flächen-orientierte Verfahren bezeichnet, betrachten die Punktmenge des Bildes als Gesamtheit. Eingesetzte Verfahren sind u.a. Region Growing [49, S. 455], Split and Merge, Pyramid Linking oder Region Merging. Alle dieser Verfahren erzeugen prinzipiell zusammenhängende Segmente. Die Überschneidungsfreiheit der Segmente ist allerdings nicht garantiert. Probleme Region-orientierter Verfahren sind u.a. die Wahl des Keim oder Seed Punktes (Region Growing), die Formulierung des Homogenitätskriteriums und Abbruchkriteriums.

Kantenbasierte Verfahren gehen auf die Idee zurück, dass Kanten häufig den Konturen von Objekten entsprechen. Gegenüber Pixel- oder Region-orientierter Verfahren werden die Segmente nicht als Flächen sondern deren Umrandungen extrahiert. Für die Extraktion eines Segments folgt auf die Kantenidentifikation meist eine Kantenverfolgung, deren angestrebtes Resultat eine zusammenhängende Kontur ist, die das Segment umrahmt.

Probleme für die Verfolgung stellen Verzweigungen, Kreuzungen sowie Lücken in der Kontur dar. Eine einmalige Fehlberechnung führt oft zu falscher Kontur.

Modellbasierte Verfahren verwenden a-priori-Wissen bzw. ein Modell, welches über das gesuchte Segment vorliegt, um dieses anhand einer parametrisierten Suche im Bild wiederzufinden. Der Einsatz einer modellbasierten Suche liefert bei bekannten Segmenten gegenüber einer generellen Segmentierung mit pixel-, region- und kantenbasierten Verfahren oft ein besseres Ergebnis. Der Einsatz von modellbasierten Ansätzen ist allerdings auf a-priori-Wissen angewiesen, was den Einsatz auf Szenarien mit bekannten Modellen beschränkt. Modellbasierte Ansätze sind u.a. Hough-Transformation, Template-Matching oder der Einsatz von statischen Modellen.

2.4.2 Kantendetektion

Überblick

Eine Kantendetektion ist eine lokale Operation um flächige, homogene Bereiche innerhalb der Bildebene voneinander zu trennen. Voraussetzung dafür ist, dass sich die Flächen hinsichtlich ihrer Grau-, Farbwerte, Helligkeit oder Textur hinreichend unterscheiden. Die Aufgabe des Detektors ist die Identifikation dieser Schwelle zwischen homogenen Bereichen und die Markierung als Kante. Darüber hinaus sollen homogene Bereiche als diese auch wahrgenommen und nicht geteilt werden.

Eine Kante ist eine Änderung der Intensität⁸ auf engstem Raum entlang einer ausgeprägten Richtung. Je stärker die Änderung ausfällt, desto intensiver wird die Struktur als Kante interpretiert.

Zur Identifikation der Intensitätsvariationen sowie der Unterdrückung von konstanten Intensitätswerten sind Ableitungsoperationen geeignet. Die Stärke einer Intensitätsänderung über einen begrenzten Ausdehnungsbereich kann mittels der ersten Ableitung über deren Verlauf bestimmt werden. Diese besitzt je nach Intensitätsveränderung ein Maximum oder ein Minimum an der Position der Kante. Mittels der zweiten Ableitung wird die Position der Kante bzw. deren steilster Anstieg oder Abfall durch einen Nulldurchgang repräsentiert. Die partiellen Ableitungen eines Bildes entlang beider Koordinatenachsen erfolgt auf Pixelebene mit,

$$\frac{\partial I}{\partial u}(u, v) \text{ bzw. } \frac{\partial I}{\partial v}(u, v), \quad (2.40)$$

⁸ Intensität bezeichnet die Energie pro Zeit pro Fläche, $[J/(sm^2)]$ bzw. $[W/m^2]$.

wobei u und v die Koordinatenrichtungen repräsentieren. Basierend auf einem zweidimensionalen Bild I ergibt sich der Gradientenvektor

$$\nabla I(u, v) = \left[\frac{\partial I}{\partial u}(u, v), \frac{\partial I}{\partial v}(u, v) \right]^T. \quad (2.41)$$

Der Betrag des Gradientenvektors,

$$|\nabla g| = \|\nabla g\|_2 = (\nabla g^T \nabla g)^{\frac{1}{2}} = \left(\sum_{w=1}^W \left(\frac{\partial g}{\partial x_w} \right)^2 \right)^{\frac{1}{2}} \quad (2.42)$$

$$|\nabla I| = \sqrt{\left(\frac{\partial g}{\partial u} \right)^2 + \left(\frac{\partial g}{\partial v} \right)^2}$$

ist gegen Rotationen des Koordinatensystems invariant [49, S. 334]. Diskrete digitale Bilder sind nicht differenzierbar weshalb das Differenzieren durch das Anwenden von diskreten Differenzen substituiert wird. Ein Vorgehen ist die Schätzung der Steigung, erste Ableitung, des entsprechenden Bildelementes $\mathbf{q}(u, v)$. Die Berechnung der Steigung eines einzelnen Bildelementes $\mathbf{q}(u, v)$ ist nicht möglich. Mittels der Bildelemente $\mathbf{q}(u-1, v)$ und $\mathbf{q}(u+1, v)$, welche $\mathbf{q}(u, v)$ beispielhaft entlang v einrahmen, lässt sich die zu $\mathbf{q}(u, v)$ gehörige Tangente und deren Steigung entlang v approximieren

$$\frac{df}{du} \approx \frac{f(u+1) - f(u-1)}{2} = \frac{1}{2} (f(u+1) - f(u-1)). \quad (2.43)$$

In der Praxis wird Gleichung 2.43 als 1D-Rechteckfilter für beide Dimensionen des Bildes I umgesetzt

$$\mathbf{H}_x = \begin{bmatrix} -0.5 & 0 & 0.5 \end{bmatrix} = 0.5 \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \text{ und} \quad (2.44)$$

$$\mathbf{H}_y = \begin{bmatrix} -0.5 & 0 & 0.5 \end{bmatrix}^T = 0.5 \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^T.$$

Die Ausrichtung der Filtermatrizen zeigt, dass H_x auf vertikale und H_y auf horizontale Kanten anspricht. Eindimensionale Filter sind anfällig für Rauschen. Eine Erweiterung auf einen 2D-Rechteckkern ermöglicht eine Glättung orthogonal zur Ableitungsrichtung. Die generierten Filterkerne

$$H_x^P = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, H_y^P = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \quad (2.45)$$

werden auch als Prewitt-Operatoren bezeichnet.

Kantendetektoren sollten in eine Richtung ableiten bzw. differenzieren, in allen anderen Richtungen allerdings glätten. Ein Differenzfilter, der zusätzlich eine glättende Komponente enthält, wird als regulierter Kantendetektor bezeichnet. Eine Gruppe von regulierenden Detektoren sind Filterkerne, die auf der Gauß-Funktion, wie bspw.

$$g_{mn} = \frac{1}{2\pi\sigma^2} e^{-\frac{n^2+m^2}{2\sigma^2}}, \quad (2.46)$$

beruhen. Der auf 2.46 beruhende Filterkern

$$\mathbf{H}^G = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (2.47)$$

ergibt in Kombination mit einem Filterkernen 2.45 einen Sobelfilter

$$\mathbf{H}_x^S = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{H}_y^S = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \quad (2.48)$$

Die Anwendung eines der vorgestellten Filterkerne auf ein Bildelement $\mathbf{I}(u, v)$ entspricht einer linearen Faltung. Das Resultat ist der angesprochene Betrag des Gradientenvektors

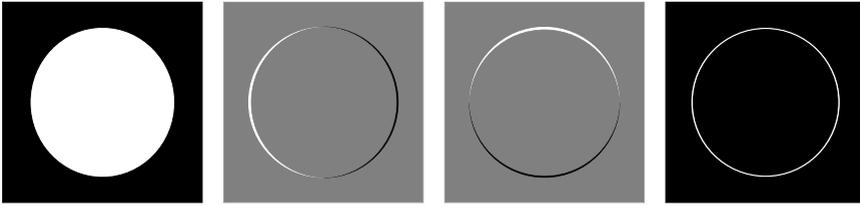
$$D_x(u, v) = \mathbf{H}_x \mathbf{I}(u, v) \quad \text{und} \quad D_y(u, v) = \mathbf{H}_y \mathbf{I}(u, v) \quad (2.49)$$

$$|\nabla I(u, v)| = \sqrt{D_x(u, v)^2 + D_y(u, v)^2}$$

Visualisiert wird die Anwendung des Sobel-Operators in Abbildung 2.12. Die Teilabbildung a) zeigt das Ausgangsbild. Die Gradientenbilder b) und c) zeigen jeweils die Anwendung der Filterkerne \mathbf{H}_x^S sowie \mathbf{H}_y^S auf das Ausgangsbild. In Binärbild d) ist das kombinierte Resultat beider Faltungen zu sehen. Aufbauend auf der isolierten Anwendung von Filterkernen wurden mehrstufige Verfahren entwickelt, um die Aussagekraft der Kantendetektoren zu erhöhen. Einer der bekanntesten mehrstufigen Ansätze ist der Canny-Algorithmus [20].

Canny-Operator

Die Anforderungen, die Canny [20] an seinen Ansatz stellte, waren, dass ausschließlich tatsächliche aber keine falschen Kanten gefunden werden,



(a) Ausgangsbild (b) Ergebnisbild durch H_x^S Anwendung. (c) Ergebnisbild durch H_y^S Anwendung. (d) Kombination der Ergebnisbilder b und c.

Abbildung 2.12: Exemplarische Anwendung eines Sobel-Operators.

dass der Abstand zwischen der gefundenen und der tatsächlichen Kante minimal ist und dass es zu keiner Mehrfachdetektion ein und derselben Kante kommt.

Der Canny-Algorithmus wird in vier Stufen unterteilt. In der Vorverarbeitung 4.3.1 wird eine Glättung des Ausgangsbildes I mittels eines Gauß-Filters 2.47 umgesetzt. Auf dem geglätteten Bild erfolgt eine Kantendetektion, wobei neben der Kantenstärke ∇I auch die Kantenrichtung

$$\Theta(u, v) = \text{actan2} \left(\frac{\partial g / \partial v}{\partial g / \partial u} \right) \quad (2.50)$$

bestimmt wird. Für die Kantendetektion werden bekannte Filterkerne wie Prewitt 2.45, Sobel 2.48 eingesetzt. Aufbauend auf den Resultaten von $\nabla I(u, v)$ und $\Theta(u, v)$ werden alle Kanten unterdrückt, die kein lokales Maximum in die Richtung des Gradienten repräsentieren. Dieses Vorgehen reduziert Mehrfachdetektionen und die eigentliche Position der Kante wird offengelegt. Durch die Anwendung werden Kanten bis auf die Breite eines Pixels reduziert. Das endgültige Pixel hat folglich lokal den größten Betrag des Gradienten. Schlussendlich folgt eine Hysterese der restlichen Kanten. Unter der Verwendung von zwei Schwellwerten T_l und T_h , wobei $T_l < T_h$ und $\frac{1}{3}T_h \leq T_l \leq \frac{1}{2}T_h$ ist, werden „schwache“ Kanten identifiziert und, je nach ihrer Nachbarschaft zu „starken“ Kanten, entfernt oder verstärkt. Das Ergebnis dieses Schrittes ist ein Binärbild, wobei ein Pixel \mathbf{q} mit

einem Gradientenvektor $h(\mathbf{q})$ wie folgt,

$$k(b) = \begin{cases} 1 & \text{if } h(\mathbf{q}) \geq T_h, \\ 1 & \text{if } T_l \geq h(\mathbf{q}) < T_h \text{ und } k(\mathbf{n}) = 1, \text{ wobei } \mathbf{n} \text{ Nachbarpixel von } \mathbf{q}, \\ 0 & \text{sonst} \end{cases} \quad (2.51)$$

auf das Binärbild abgebildet wird.

Der Canny-Algorithmus gilt zurzeit weithin als optimal unter den klassischen Kanten-Operatoren bzgl. Erkennung⁹, Lokalisierung¹⁰ und Ansprechverhalten¹¹. Die Ergebnisse des Canny-Filters hängen allerdings sehr stark von der Hysterese bzw. den Parametern T_l und T_h ab. Diese werden weithin manuell über mehrere Testläufe iterativ bestimmt, um ein optimales Ergebnis zu erzielen. Eine manuelle Festlegung der Parameter ist bei einer automatisierten Anwendung unter dynamischen Umwelt- und Beleuchtungsbedingungen auf wenige Szenarien beschränkt. Eine dessen ungeachtete Verwendung fester Parameter führt zu Fehlinterpretationen der Kanten.

Histogramm-Konkavitätsanalyse

Eine dynamische Schätzung der Parameter T_l und T_h ohne manuellen Eingriff wird in der Arbeit von Rosenfeld und La Torre [76] mittels einer Histogramm-Konkavitätsanalyse vorgestellt. Dazu wird als Vorverarbeitung das Histogramm $h(g)$ des Ausgangsbildes untersucht, um auf dessen Basis die Schwellwerte T_h und daraus T_l zu bestimmen. Zur Herleitung des Schwellwertes aus dem Histogramm, wird das Histogramm mit einer konvexen Hülle $\bar{h}(g)$ überzogen und die Differenz

$$k(g) = \bar{h}(g) - h(g) \quad (2.52)$$

zwischen der Hülle und dem Histogramm bestimmt, wobei $g_s < g < g_e$ und g_s den kleinsten und g_e den größten Grauwert ungleich Null repräsentieren. Die beschriebene Situation wird in Abbildung 2.13 beispielhaft skizziert. Der gesuchter Schwellwert T_h entspricht damit dem Wert g , für welchen gilt $\max \{k(g)\}$.

Als Maß der Konfidenz für $\max \{k(g)\}$ setzen Rosenfeld und La Torre [76] die Position von g innerhalb der Verteilung des Histogramms an. Dieses

⁹Alle tatsächlichen Kanten werden gefunden, aber auch keine falschen.

¹⁰Abstand zwischen tatsächlicher und erkannter Kante ist möglichst klein.

¹¹Keine Mehrfacherkennung ein- und derselben Kante.

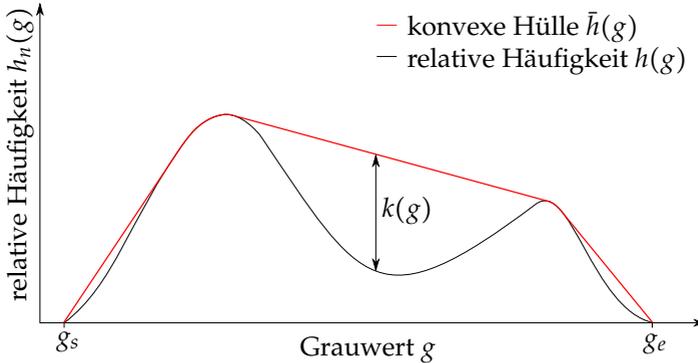


Abbildung 2.13: Exemplarisches Histogramm mit konvexer Hülle zur Herleitung der Canny-Operator Hystereseparameter.

soll verhindern, dass $\max \{k(g)\}$ auf einer nur scheinbar konvexen Hülle beruht. Eine scheinbar konvexe Hülle würde sich ergeben, wenn sich das gefundene Maximum am Rand des Histogramms befindet und sich damit der Großteil des Histogramms einseitig auf einer Seite des Maximums befindet. Einseitige Verteilungen des Histogramms gegenüber $\max \{k(g)\}$ sollen verhindert werden, da es zur Unterdrückung von starken Kanten bzw. Stärkung von schwachen Kanten führt. Die Position von g innerhalb der Verteilung des Histogramms wird mittels des Verhältnisses der relativen Häufigkeiten

$$E(g) = \left(\sum_{j=g_s}^{g-1} h(j) \right) \left(\sum_{i=g+1}^{g_e} h(i) \right), \quad (2.53)$$

die sich auf beiden Seiten von g befinden, bestimmt. Entspricht $g = (g_s \vee g_e)$ folgt $E(g) = 0$. Auf der anderen Seite ist $E(g)$ maximal, wenn g dem Mittelwert des Histogramms entspricht. Resultierend soll ein T_h ignoriert werden, das auf einem g mit einem geringen $E(g)$ beruht.

2.5 Hindernisrepräsentation

2.5.1 Überblick

Für die Fusion von Sensordaten im Bereich der Hindernisdetektion und -vermeidung werden oftmals approximierende Abbildungen der Umwelt verwendet [6, 44, 80]. Bahnplanungsansätze [2] nutzen als Grundlage des Öfteren ebenfalls ein Weltmodell, welches Position und Abmessungen von

bekannten Hindernissen und weitere Informationen, die für die Planung Relevanz aufweisen, enthält. Das Weltmodell, welches auch als Hindernisrepräsentation oder -karte bezeichnet wird, bietet sich als Schnittstelle an, um Hindernisinformationen sowie Aktualisierungen zwischen dem Sensor und der Pfadplanung auszutauschen.

Bei einer Karte mit dem Ziel, Hindernisse zu repräsentieren, handelt es sich im Allgemeinen, um ein vereinfachtes Modell der Umwelt. Ziel einer Hindernisdetektion ist nicht die Identifikation von Objekten oder die Klassifizierung nach Baum, Haus oder Objekt XY. Vielmehr wird eine Detektion von freien und belegten Räumen angestrebt. Belegte Räume enthalten Hindernisse und sind nicht ohne die Gefahr einer Kollision von einem Vehikel zu durchqueren. Den Grenzflächen zwischen belegten und freien Räumen können aber durchaus weitere Eigenschaften zugeordnet werden, etwa ob sie zum Landen geeignet sind.

Die Begrifflichkeiten Weltmodell, Hindernisrepräsentation oder Karte werden im Rahmen der Arbeit für Karten verwendet, die Objekte oder Gefahren repräsentieren. Objekte bzw. Gefahren belegen Raum, welcher von den Fahrzeugen folglich nicht durchquert werden kann. Die Belegung von Raum wird mittels Sensordaten detektiert. Die Sensordaten zeigen im Allgemeinen pro Messvorgang nur eine Teilabbildung der Umwelt. Einen umfassenderen Überblick über mögliche Gefahrenquellen, die den Sensorträger bzw. den Sensor umgeben, wird erst durch die Kombination mehrerer Sensormessungen erreicht. Die räumliche und zeitliche Kombination mehrerer Sensordaten in einem Weltmodell wird als Kartierung bezeichnet. Weltmodelle können in metrische, topologische und semantische Karten unterteilt werden.

Metrische Karten präsentieren Objekteigenschaften wie Abmessungen, Form und Position mittels geometrischer Elemente wie Punkten, Geraden, Flächen, dreidimensionale Körper oder einer diskreten Rasterung der Umgebung.

Topologische Karten sind Graphen, deren Knoten wichtige Punkte und deren Kanten die Verbindungswege zwischen den Knoten repräsentieren. Graphen können zur Pfadplanung eingesetzt werden, bspw. für Straßennetze zwischen Städten oder Verbindungen zwischen Straßenkreuzungen einer Stadt.

Semantische Karten können sowohl durch metrische als auch durch topologische Karten repräsentiert werden, enthalten aber noch zusätzliche Informationen, die über die der zugrunde liegenden Karten hinausgehen.

Semantische Karten dienen nicht im eigentlichen Sinne dem Kartographieren von Hindernissen. Es werden vielmehr ausgewählte Objekte innerhalb der Karte mit zusätzlichen Informationen verknüpft. Darunter fallen bspw. Informationen, ob eine Oberfläche zum Landen geeignet ist.

Je nach Zielstellung werden in der Literatur Präferenzen für die unterschiedlichen Kartentypen ausgedrückt. Metrische Karten weisen Vorteile bei der Sensordatenfusion auf [7, 46], wogegen topologische Karten in der globalen Pfadplanung bevorzugt werden.

Es werden auch hierarchische Unterschiede verdeutlicht. Diskrete Rasterkarten sind gut für die Fusion von Daten geeignet. Körper oder deren Oberflächen, welche auf kombinierten Rasterelementen aufsetzen, sind dagegen besser für Kollisionstests geeignet, die in der Pfadplanung Einsatz finden.

Die jeweiligen Vor- und Nachteile der einzelnen Kartentypen führten in Anwendungen, in denen mehrere Parteien an einem Weltmodell substituieren zu einer Kombination der verschiedenen Kartentypen. Diese wird durch einen Informationsaustausch zwischen den einzelnen Kartentypen, wie bei Andert [7] gezeigt, erreicht.

2.5.2 Belegtheitsgitter

Belegtheitsgitter sind metrische Weltmodelle, welche den Raum kubisch diskretisieren [68]. Die Orientierung des Weltmodells ist identisch mit dem Koordinatensystem, in dem sich die einzutragenden Objekte befinden. Jede kubische Einheit, im Folgenden auch als Zelle bezeichnet, enthält Informationen über seine Belegung. Die Belegung ist ein Wahrscheinlichkeitswert, welcher Auskunft über die Anwesenheit eines Objektes innerhalb des durch die Zelle beschriebenen Raums gibt.

Zellen sind Teil eines Feldes, welches eine lückenlose Rasterung der abzubilden Umwelt erlaubt. Bezeichnet werden diese Felder unter anderem als Belegtheitsgitter M (engl.: certainty, evidence oder occupancy grid) [67].

Unter der Annahme, dass die Sensoren ideale Daten bereitstellen, wird jede Zelle ausschließlich mit den Zuständen „frei“ oder „belegt“ kategorisiert. Für die inkrementelle Erzeugung eines Belegtheitsgitters M wird von einem „freien“ Belegtheitsgitter $M = \emptyset$, in dem keinerlei Objekte enthalten sind, ausgegangen. Die Messung eines Sensors, dessen Pose im Raum bekannt ist, ist eine Entfernungsmessung zu einer „undurchdringlichen“ Objektoberfläche. Die generierte Messung wird als Objektpunkt \mathbf{p} aufgefasst, der den Schnittpunkt zwischen Objektoberfläche und Messstrahl beschreibt. Dieser wird vom Sensorkoordinatensystem in das Koordinaten-

system von M überführt. Die Zelle, welche den jeweiligen \mathbf{p} umrahmt, wird als „belegt“ gekennzeichnet, da sie unabhängig vom tatsächlich angewendeten Messverfahren als undurchdringlich angenommen wird. Mehrere Messungen aus unterschiedlichen Posen oder zu unterschiedlichen Zeitpunkten ermöglichen Aussagen über eine Vielzahl von Zellen. Das Resultat ist ein Belegtheitsgitter mit freien und belegten Zellen, welches, wie eine Karte, Auskunft über Position und Ausmaße der vermessenen Objekte gibt.

Eine ideale Sensormessung, der bedingungslos vertraut werden kann, ist auf Anwendungen mit realen Sensoren, welche Ungenauigkeiten, Fehler und Unsicherheiten enthalten, wie in Abschnitt 2.3.2 dargelegt, nur bedingt übertragbar. Erweiterungen der grundsätzlichen Idee des Belegtheitsgitters zur Berücksichtigung von Sensorunzulänglichkeiten werden u.a. von Elfes [23] und Raschke und Borenstein [74] für das Sonar, von Andert [7] und Hrabar [46] für die Stereokamera und von Scherer [79] für Laserscanner vorgestellt.

Die Berücksichtigung von Sensorunzulänglichkeiten führte zu einer Erweiterung des Wertebereiches jeder Zelle, um ebenso unsichere Belegungen abzubilden. Borenstein [16] verwendet als Indikator für die Signifikanz der Belegtheit die Anzahl der Messungen, welche die Zelle als Belegung markieren. Wiederholt als belegt eingeordnete Zellen deuten auf Basis unabhängiger, unkorrelierter Messungen auf die Anwesenheit von realen Objekten hin. Vereinzelt Messungen lassen auf Rauschen schließen.

Murray [69] berücksichtigt zusätzlich den freien Raum zwischen dem bestimmten Objektpunkt \mathbf{p} und der Sensorquelle als frei. Diese Betrachtung ermöglicht es, widersprüchliche Aussage zu entkräften und gleichorientierte Aussagen weiterhin zu verstärken.

Eine Auswertung der Häufigkeit, mit der eine Zelle als belegt deklariert wurde, setzt identische Unsicherheiten aller eingesetzten Messungen voraus. Eine Kombination von unterschiedlichen Unsicherheiten, auf Basis von Daten unterschiedlicher Sensoren oder auch innerhalb eines Datensatzes, würde durch eine alleinige Auswertung der Häufigkeit nicht korrekt interpretiert werden.

Gelöst wurde diese Problematik mittels des Ansatzes der Belegtheitswahrscheinlichkeit (engl.: occupancy probability grid), präsentiert u.a. von Elfes [24]. Jeder Sensormessung wird je nach Sensoreigenschaften eine Belegtheitswahrscheinlichkeit zugeordnet, welche mit der Belegtheitswahrscheinlichkeit jeder durchdrungenen bzw. berührten Zelle kombiniert wird. Neben der Fusion von Sensordaten erlaubt der Einsatz von stochastischen Methoden die Repräsentation von unbekanntem Gebieten innerhalb eines Belegtheitsgitters M .

Notation

Die Abbildung einer Koordinate $\mathbf{p}_i : (x_i, y_i, z_i)^T$ im System i in das Belegtheitsgitter M , welches kubusförmigen Zellen aufweist, erfolgt mittels

$$\mathbf{p}_m = \left\lfloor \frac{\mathbf{p}_i - \mathbf{t}_{mi}}{d_m} + 0.5 \right\rfloor. \quad (2.54)$$

Die Koordinate \mathbf{p}_m beschreibt das Zentrum einer Zelle m , deren Kantenlänge d_m ist. Die Diskretisierung 2.54 einer kontinuierlichen Umwelt auf die Maße des Belegtheitsgitters ist verlustbehaftet.

Der Ursprung von M liegt in der Zelle m mit der Koordinate

$\mathbf{p}_m : (x_m, y_m, z_m)^T = (0, 0, 0)^T$, welche bezüglich des Ursprungs des Systems i um den \mathbf{t}_{mi} translatorisch verschoben ist. Die Rücktransformation von \mathbf{p}_m in die kontinuierlichen Koordinaten des Systems i ist

$$\mathbf{p}_i = \mathbf{p}_m \cdot d_m + \mathbf{t}_{mi}. \quad (2.55)$$

Jede Zelle $m \in M$ beschreibt die Belegung des von ihr umfassenden Raums binär

$$M : o_m(x_m, y_m, z_m) \mapsto \{0, 1\}, \quad (2.56)$$

wobei $o(x_m, y_m, z_m) = 1$ die Belegung der betrachteten Zelle durch ein Objekt verdeutlicht. Die Berücksichtigung der Unsicherheiten einer Belegung von m durch ein Objekt erfolgt mittels

$$p_m(x_m, y_m, z_m) \rightarrow [0, 1]_{\mathfrak{R}}. \quad (2.57)$$

Um bei der Fusion von Sensormessungen für ein m die Unterscheidbarkeit von Belegungswahrscheinlichkeiten nahe den Grenzen 0 und 1 zu gewährleisten, wird u.a. nach Thrun [88, S. 286] die Belegungswahrscheinlichkeit p in einen Logitwert l überführt, wobei

$$l_m = \ln \frac{p_m}{1 - p_m}, l \in \mathfrak{R} \quad (2.58)$$

entspricht. Eine Abbildung der Logitwerte und der Belegungswahrscheinlichkeiten auf die binäre Repräsentation von m für jede Zelle erfolgt mittels

$$o_m = \begin{cases} 1 & \text{für } p_m > p_{unknown} \text{ bzw. } l_m > l_{unknown}, \\ 0 & \text{sonst.} \end{cases} \quad (2.59)$$

Ein Ausgangspunkt für den Schwellwert der Operation ist $p_{unknown} = 0,5$ bzw. $l_{unknown} = 0$. Da die tatsächliche Anwesenheit eines Objektes an der

Position der jeweiligen Zelle mit zunehmender Wahrscheinlichkeit $p(Z, Z_p)$ korreliert, würde eine höhere Schwelle die Identifikation von möglichen Phantomobjekten, die durch Rauschen entstehen, reduzieren. Gleichzeitig wäre aber die Gefahr größer, dass tatsächliche Objekte "übersehen" werden würden. Eine Reduktion $p_{unknown}$ würde dementsprechend zu einer verstärkten Berücksichtigung von Rauschen bei der Kartierung führen.

Inkrementelles Update

Unter Verwendung der Transformation $({}^i T_s)_t$ aus Abschnitt 2.2.3 und einem inversen Sensormodell (2.3.2) werden die Sensorposition und aufgezeichneten Messungen des Zeitpunktes t inkrementell aus dem Sensor- in das Weltkoordinatensystem und schlussendlich in die Koordinaten des Belegtheitsgitters (2.54) überführt.

Unter Kenntnis der Sensorposition $\mathbf{O}_{s,t}$ und einer Approximation \mathbf{p}_m zum Zeitpunkt t im Belegtheitsgitter wird ein Strahl bestimmt, der ausgehend von $\mathbf{O}_{s,t}$ durch \mathbf{p}_m verläuft. Die Länge des Strahls ist auf die maximale Sensorreichweite begrenzt.

Zur Identifikation der Zellen, die vom Strahl geschnitten werden, wird diesem, ausgehend von $\mathbf{O}_{s,t}$, gefolgt. Ein gängiges Verfahren, welches die Darstellung einer Linie in einer gerasterten zweidimensionalen Umgebung ermöglicht, wurde von Bresenham [17] präsentiert. Eine Erweiterung des Verfahrens von Bresenham auf ein dreidimensionales Anwendungsfeld wird von Andert [7, S. 45] aufgezeigt. Ebenfalls ist eine Nutzung von Ray-Tracing Ansätzen [5], [8, Kapitel 12.2] möglich.

Für jede der identifizierten Zellen wird, bedingt durch ihren kartesischen Abstand Z zu $\mathbf{O}_{s,t}$, eine Belegungswahrscheinlichkeit $p(Z, Z_p)$ auf Basis des inversen Sensormodells 2.3.2 bestimmt. Das Resultat $p(Z, Z_p)$ für den Zeitpunkt t wird mit dem Wert, den die Zelle bis zum Zeitpunkt $t - 1$ vertrat, kombiniert. Die Kombination der Belegtheitswahrscheinlichkeit p einer Zelle m aus dem Zeitpunkt $t - 1$ und dem Zeitpunkt t wird auf Basis des Satz von Bayes

$$\frac{p_{m,1:t}}{1 - p_{m,1:t}} = \frac{p_{m,1:t-1}}{1 - p_{m,1:t-1}} \cdot \frac{p_{m,t}}{1 - p_{m,t}} \cdot \frac{1 - p_{m,0}}{p_{m,0}} \quad (2.60)$$

bestimmt [85, S. 36]. Durch eine Substitution der Belegungswahrscheinlichkeiten mit Logit-Werten ergibt sich

$$l_{m,1:t} = l_{m,1:t-1} + l_{m,t} - l_{m,0}. \quad (2.61)$$

Voraussetzung für dieses Vorgehen ist eine Initialisierung jeder Zelle m zum Zeitpunkt $t = 0$ ($p_{m,0}$ bzw. $l_{m,0}$) mit einem definierten Zustand. Liegt

zu $t = 0$ kein Wissen über die Umgebung vor, ist eine Initialisierung mit $p_{m,0} = p_{unknown}$ bzw. $l_{m,0} = l_{unknown}$ denkbar. Visualisiert wird das inkrementelle Einfügen von Messungen \mathbf{p} am Beispiel eines zweidimensionalen Belegtheitsgitters in Abbildung 2.14. In der Abbildung repräsentieren weiße Zellen unbekannte Gebiete, rote Zellen entsprechen Objekten und grüne Zellen sind freie Bereiche. Ausgehend von einer Messung in einem vollkommen unbekanntem Gebiet (2.14a) werden fünf Messungen inkrementell hinzugefügt (2.14b). Beim Hinzufügen der Messungen überschneiden sich mehrere Strahlen in einigen Zellen. Im Beispiel weisen die Strahlen an der Position der geschnittenen Zellen einen Wert $p(Z, Z_p) < p_{unknown}$ auf. Die Kombination der $p(Z, Z_p)$ pro Zelle mit dem Satz von Bayes führt zu einer Bestätigung der Freizeichnung. Visualisiert ist die wiederholte Bestätigung einer Belegung (frei oder belegt) durch eine Kräftigung des Farbwertes. Am Beispiel der Freizeichnung wird aus einem blassen Grün bei einmaliger Freizeichnung ein kräftiges Grün bei einer mehrmaligen Bestätigung. Die Stärkung der Signifikanz, die in Abbildung 2.14 für das Frei-

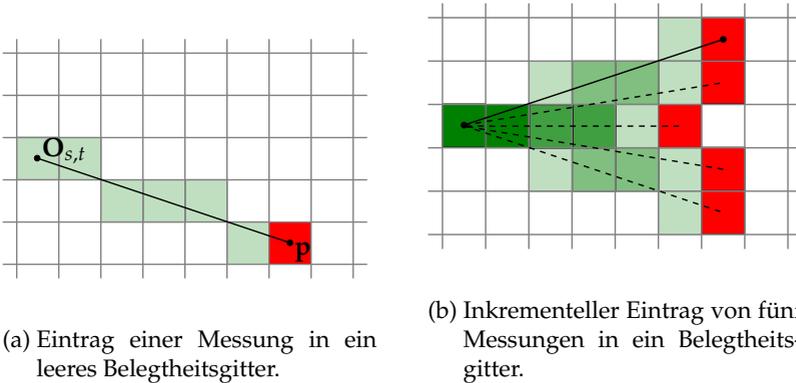


Abbildung 2.14: Zweidimensionales Belegtheitsgitter, in das inkrementell Messungen \mathbf{p} eingefügt werden. Nach Scholz [81].

zeichnen von Zellen gezeigt wurde, ist natürlich auf eine Bestätigung der Belegung von belegten Zellen übertragbar.

2.5.3 Datenstruktur

Belegtheitsgitter können mittels Feldern (engl. array) [7, 44, 79] oder Bäumen [63, 72, 98] aufgebaut werden. Felder weisen gegenüber Bäumen allerdings den grundlegenden Nachteil des wesentlich höheren Speicherbedarfs auf. Bei der Initialisierung eines Feldes ist es notwendig, dass jede

Zelle initialisiert wird. Das führt dazu, dass nach der Initialisierung das gesamte Feld im Speicher abgebildet ist, unabhängig von der tatsächlichen Verteilung und Anzahl der belegten Zellen innerhalb des Feldes. Dadurch ist in größeren Freiluft-Kartierungen oder bei angestrebter feiner Auflösung der Kartierung das benötigte Speichervolumen nicht mehr tragbar [98].

Im Folgenden wird ein grundlegender Überblick über die Datenstruktur der Bäume gegeben, bevor auf die verwendete Baumstruktur OcTree eingegangen wird. Dabei wird vergleichend immer wieder auf Felder eingegangen.

Baumstruktur

Im Sinne einer Datenstruktur wird ein Baum als ein azyklischer, zusammenhängender Graph bezeichnet, in dem Knoten durch Kanten verbunden sind. Abhängig davon, ob die Kanten des Baumes eine spezifizierte Richtung aufweisen, können Bäume in ungerichtete und gerichtete Bäume unterteilt werden. Eine enger gefasste Bezeichnung für den gerichteten Baum ist der gewurzelte Baum, wobei ein Knoten, die Wurzel definiert, welche der Ausgangspunkt des Baumes ist. Gewurzelte Bäume erlauben die Umsetzung einer Hierarchie, wobei eine Unterteilung in Eltern- und Kind-Knoten üblich ist. Kind-Knoten sind immer auf einer ihren Eltern untergeordneten Ebene angeordnet. Knoten, die auf einem identischen Eltern-Knoten basieren, werden Geschwister genannt. Knoten ohne Kinder werden als Blätter bezeichnet. Knoten, die keine Blätter sind, sind innere Knoten. Die Länge eines Pfades von der Wurzel zu einem beliebigen Knoten ist die Tiefe d dieses Knotens. Die maximale Tiefe, die ein Knoten in einem Baum erreichen kann, wird als Höhe h des Baumes bezeichnet. Abbildung

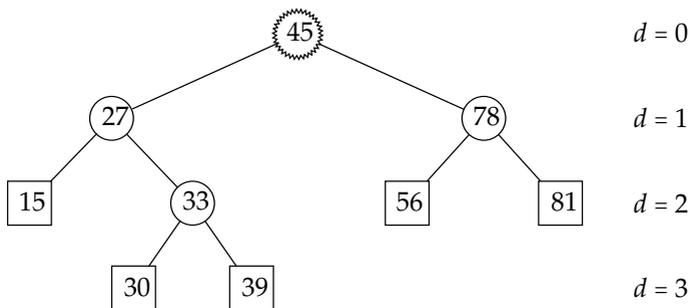


Abbildung 2.15: Geordneter, gewurzelter Binärbaum¹² der Höhe 3.

2.15 verdeutlicht am Beispiel eines Binärbaums die oben beschriebenen Termini. Die Wurzel hat den Wert „45“. Alle inneren Knoten sind durch Kreise und alle Blätter durch Quadrate gekennzeichnet. Die Höhe des Baumes ist drei und die Tiefe bspw. des Knotens „81“ ist zwei.

Weitere Unterscheidungsmerkmale zwischen Bäumen sind, ob eine Ordnung innerhalb des Baumes vorliegt oder ob der Baum vollständig ist. Eine Ordnung innerhalb des Baumes liegt vor, wenn unter den Kindern eines Knotens eine Ordnung festgelegt ist, etwa über eine „größer“ Relation, wie in Abbildung 2.15. Bäume ohne eine interne Sortierung werden als ungeordnet bezeichnet.

In einem vollständigen Baum weisen alle Knoten die maximale Anzahl an Kindern auf und alle Blätter haben die gleiche Tiefe.

Darüber hinaus können Bäume über die Art der Datenhinterlegung unterschieden werden. Es wird zwischen blattorientierter und knotenorientierter Datenhinterlegung unterschieden. Bei der blattorientierten Hinterlegung enthalten nur die Blätter die Zieldaten und bei der knotenorientierten nur die inneren Knoten [14, S. 10].

OcTree

Ein OcTree (octo lat. acht, tree engl. Baum) ist ein gewurzelter, geordneter Baum, dessen Knoten jeweils acht Kinder oder keines aufweisen. OcTree erlaubt eine geordnete hierarchische Abbildung von dreidimensionalen Daten und damit auch die Repräsentation von kubischen Koordinaten. Die Verwendung eines Knotens nicht ausschließlich als Repräsentant einer einzelnen Koordinate, sondern als Repräsentation eines kubischen Raumes, erlaubt es, mittels eines OcTree eine rekursive Dekomposition des Raumes zu gleichen Teilen vorzunehmen. Eine Umsetzung dieses Konzepts wurde unter anderem von Meagher [8, 63, 72, 98] vorgestellt. Ein Beispiel einer derartigen Zerlegung eines Raumes zeigt Abbildung 2.16. Die Wurzel eines OcTree weist ebenfalls wie jeder Knoten im OcTree die Form eines Kubus (rot) auf und beinhaltet den kompletten zu beschreibenden Raum. Kind-Knoten besitzen die halbe Kantenlänge ihrer Eltern-Knoten und beschreiben jeweils ein Achtel des vom Eltern-Knoten beanspruchten Volumens. Kind-Knoten werden auch als Oktanten bezeichnet. Die rekursive, hierarchische Demontage des Raumes ist prinzipiell unbegrenzt und erlaubt die Darstellung beliebiger dreidimensionaler Objekte in jeder gewünschten Auflösung [63]. Ausgehend von der Lage und Größe der Wurzel kann jedem Knoten eindeutig ein physikalischer Raum zugeordnet werden.

Der hierarchische Aufbau des OcTree erlaubt einen bedarfsabhängigen Zu-

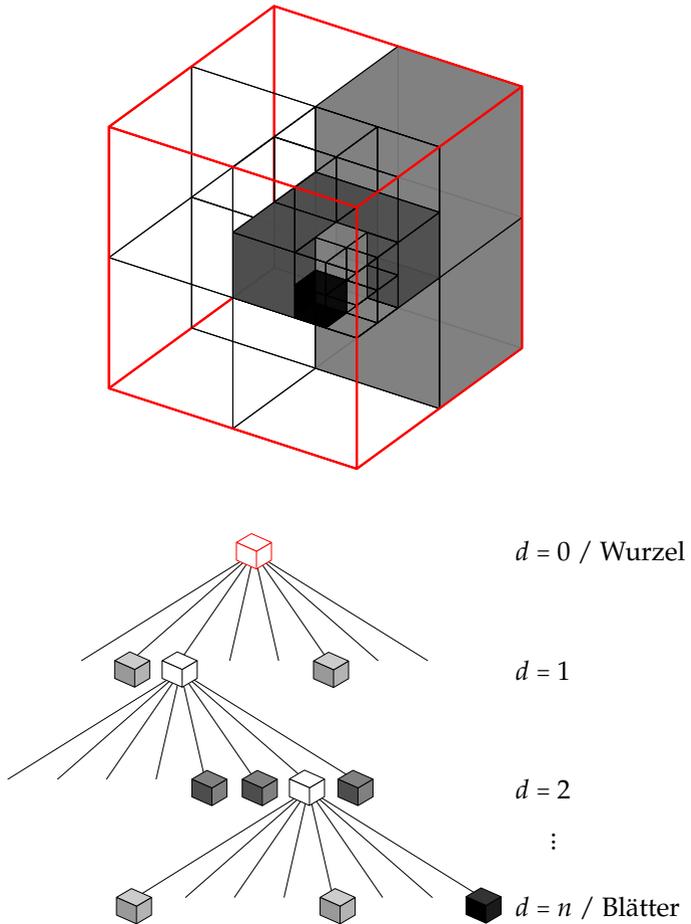


Abbildung 2.16: Rekursive Dekomposition eines dreidimensionalen Raumes mittels eines Octree der Höhe drei.

griff auf unterschiedliche Auflösungen des betrachteten Raums und ermöglicht die Manipulation, Analyse und Darstellung je nach Anforderungen auch auf einfacher Hardware.

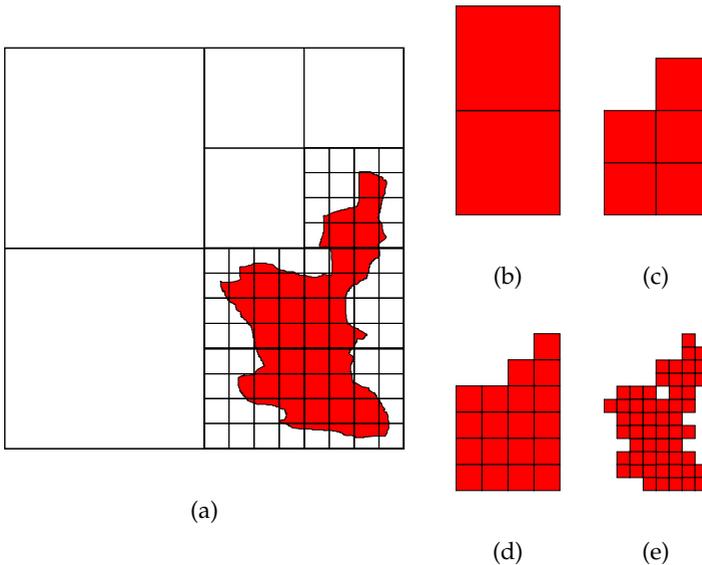


Abbildung 2.17: Rekursive Dekomposition eines zweidimensionalen Objekts mittels eines QuadTree¹³ der Höhe vier.

Die Fähigkeit des bedarfsabhängigen Zugriffs auf unterschiedliche Auflösungen erfordert allerdings, dass die Zustände der inneren Knoten auf denen der Blätter aufbauen, was eine reine Einordnung in blattorientierte und knotenorientierte Bäume nicht ermöglicht. Vielmehr muss eine Kombination aus blattorientierter und knotenorientierter Definition umgesetzt werden, so dass Zustände sowohl in Blättern als auch inneren Knoten hinterlegt sind. Kraetzschmar [53] und Wurm [98] präsentieren Varianten, die Belegtheitswahrscheinlichkeit der acht Kinder n_i , je nach Anwendungszweck auf die Eltern n zu übertragen. Eine Variante ist es, die durchschnittliche Belegtheit $L(n_i)$ der Kinder

$$\bar{L}(n) = \frac{1}{8} \sum_{i=1}^8 L(n_i) \quad (2.62)$$

¹³Baumstruktur, deren innere Knoten jeweils vier Kinder aufweisen. Besitzt ein vergleichbares Verhalten wie ein OcTree, allerdings zur Organisation von zweidimensionalen Daten.

auf den Eltern-Knoten anzuwenden. Eine zweite Umsetzung ist die Nutzung der maximalen Belegung der Kinder

$$\hat{l}(n) = \max_i L(n_i). \quad (2.63)$$

Die Nutzung der maximalen Belegung der Kind-Knoten ist die konservativere Variante, welche nach Wurm [98] gut geeignet ist, um in der Navigation von unbemannten Fahrzeugen eingesetzt zu werden. Zur Verdeutlichung, wie die Belegung der Kind-Knoten auf die jeweiligen Eltern wirkt, zeigt Abbildung 2.17 die Anwendung von 2.63 auf einen QuadTree. Von den Blättern (e) mit feinsten Auflösung bis zu (b) ist eine Reduktion der Auflösung und damit eine konservative Abschätzung der belegten Flächen zu erkennen. Die Wahl der Auflösung erlaubt es je nach Anwendung zwischen einer hohen Auflösung und damit einer hohen Präzision bei der Abbildung oder einer geringeren Auflösung bei verbesserter Berechenbarkeit zu wählen.

Kapitel 3

Voraussetzungen für multisensorielle Hindernisdetektion

Dieses Kapitel führt in zwei Themen ein, welche fundamental für die Umsetzung der Sichtfelderweiterung sind.

Abschnitt 3.1 geht auf die Ermittlung der extrinsischen Parameter zwischen monokularer Kamera und monogonem LiDAR mittels einer automatischen Kalibrierung ein. Das Augenmerk liegt verstärkt auf der Verwendung von LiDAR-Intensitätsmessungen für die Kalibrierung und die Identifikation von Fehlermessungen durch eine kombinierte Nutzung von Intensitäts- und Entfernungsmessungen.

Abschnitt 3.2 setzt sich mit der Überführung einer kontinuierlichen Welt in eine diskrete, probabilistische Umweltprepräsentation auseinander. Im Fokus der Betrachtung steht das „Übersehen“ von Hindernisstrukturen bzw. das Entfernen von zuvor gesichteten Hindernissen durch widersprüchliche Einträge in die Repräsentation, wobei die Widersprüche auf die Diskretisierung zurückzuführen sind.

3.1 Extrinsische Kalibrierung

Eine Sensordatenfusion von LiDAR- und Kameradaten auf Feature- und Pixelebene, wie sie bei der Sichtfelderweiterung angestrebt wird, setzt die Identifikation von korrespondierenden Messungen bzw. Pixeln in den Datensätzen voraus. Die Korrespondenz wird im Allgemeinen sowohl örtlich als auch zeitlich angestrebt.

Voraussetzung für die Identifikation von korrespondierenden Messungen ist wiederum die Kenntnis der jeweiligen Sensorposen in einem gemeinsamen Koordinatensystem relativ zueinander. Für die Herleitung der relativen räumlichen Zuordnung bzw. der extrinsischen Transformation cT_l

(siehe: 2.1) zwischen Kamera c und LiDAR l ist es erforderlich, die Rotation $\Phi_{cl}(\Theta_x, \Theta_y, \Theta_z)^T$ und die Translation $\mathbf{t}_{cl}(t_x, t_y, t_z)$ zwischen beiden Sensoren zu bestimmen.

Im rudimentären Fall kann auf konstruktives Vorwissen zurückgegriffen werden um bspw. die Anbaupose der einzelnen Sensoren am Sensorträger zu ermitteln. Dieses Vorgehen ist nur bedingt zu empfehlen, da zwar die Pose der Anbauposition des Sensors am Träger, aber nicht die Sensorquelle bzw. -senke, die sich innerhalb des Sensorgehäuses befinden, bekannt ist. Die Pose der eigentlichen Sensoren innerhalb des jeweiligen Gehäuses ist aufgrund von Fertigungstoleranzen oft nicht hinreichend bekannt. Eine Abschätzung der extrinsischen Parameter ergänzend zur Berücksichtigung von konstruktiven Maßen ist somit zweckmäßig.

Ein oft verwendeter Ansatz ist die Nutzung der vom Sensor selbst produzierten Messungen, um die extrinsischen und intrinsischen Parameter eines Sensors abzuschätzen. Eine hinreichend genaue intrinsische Kalibrierung ist die Voraussetzung für die Schätzung der extrinsischen Parameter. Eine intrinsische Kalibrierung für Kameras wurde u.a. von Zang [100] und für LiDAR von Okubo [71] präsentiert. Sie werden im Folgenden als hinreichend angenommen.

Die Nutzung der vom Sensor produzierten Messungen setzt voraus, dass die vom LiDAR bereitgestellten Daten die korrekte Position und Abmessungen des Kalibrierobjekts enthalten. Diese Annahme kann allerdings aufgrund von Detektionsverhalten wie Kantentreffern (siehe Abschnitt 2.3.2) nicht aufrecht erhalten werden. Vielmehr muss davon ausgegangen werden, dass durch Kantentreffer in Verbindung mit einer „First-Echo“-Auswertung betrachtete Objekte in den LiDAR-Daten größer erscheinen als sie in Realität sind [54]. Dieses „Aufblähen“ verhindert eine korrekte Vermessung des Kalibrierobjekts und hat folglich Einfluss auf die Ermittlung der extrinsischen Parameter.

Um die Einflüsse der Kantentreffer auf die Vermessung des Kalibrierobjekts zu minimieren, wird ein Ansatz zur extrinsischen Parameteridentifikation vorgestellt, der neben den Entfernungsmessungen des LiDAR verwendet. Die Intensitätsmessungen werden als Indikator für das Auftreten von Kantentreffern eingesetzt und erlauben eine genauere Definition des Kalibrierobjekts als bei einer alleinigen Verwendung von Entfernungsmessungen.

Einleitend wird ein Überblick über bestehende Ansätze zur Konkretisierung der extrinsischen Parameter gegeben. Aufbauend wird der Ansatz von Li et al. [59] skizziert, der den Kern des hier vorgestellten Ansatzes bildet. Li et al. setzen allerdings voraus, dass die generierten Informationen des LiDAR korrekt sind. Dass diese Annahme nicht garantiert werden

kann, wird anhand eines Beispiels gezeigt, in dem der Einfluss von Kantentreffern auf die Vermessung eines Objektes sowie die Möglichkeit der Identifizierung verdeutlicht wird. Um den Einfluss der Kantentreffer zu minimieren, wird ein neues Kalibrierobjekt vorgestellt, welches die Lokalisation von Kantentreffern mit den Intensitätsmessungen des LiDAR und die automatische Lokalisierung in den Kamera- als auch LiDAR-Daten erlaubt. Beide Punkte d.h. sowohl die automatische Ermittlung der Daten als auch die Identifikation der Kantentreffer führen, zu einer Steigerung der Kalibriergenauigkeit, was abschließend in einer Validierung vergleichend zu anderen Ansätzen gezeigt wird.

3.1.1 Überblick

Die Abschätzung der relativen Transformation zwischen Kamera und LiDAR, deren Lage und Position zueinander statisch ist, wird in zahlreichen Publikationen verdeutlicht. Alle Ansätze haben grundsätzlich das Ziel, in den jeweiligen Sensordaten korrespondierende Features zu identifizieren, um darauf aufbauend die Relation zwischen den Sensoren zu bestimmen. Um welche Features es sich handelt, hängt zumeist von den Eigenschaften der verwendeten Sensoren ab. Unterschieden werden können Ansätze, die mit Laserlicht im sichtbaren Spektrum arbeiten, welches wiederum von der Kamera detektiert werden kann [13, 30, 64, 77], dreidimensionale Feature-Detektion mittels 3D LiDAR und Stereokamera [4, 27, 78] und Verfahren, die monokulare Kameras und LiDAR, welche Licht emittieren, das im Kamerabild nicht detektiert werden kann [59, 70, 99], verwenden.

Das LiDAR, welches der vorgestellten Arbeit zugrunde liegt, emittiert, ebenso wie ein Großteil der heute eingesetzten LiDAR [37, 83, 93], welche in einer Kombination mit passiven Kameras außerhalb von Gebäuden eingesetzt werden, Licht im nahen Infrarotspektrum. Im Folgenden werden daher Ansätze betrachtet, welche monokulare Kameras in Verbindung mit LiDAR verwenden, wobei das emittierte Laserlicht Teil des nahen Infrarotspektrums ist, welches von den Kameras nicht detektiert werden kann.

Die Ausgangssituation ist, dass die Kamera keine Strahlung emittiert. Eine Detektion durch das LiDAR ist daher nicht möglich. Aufgrund einer ausschließlichen Detektion des sichtbaren Spektrums durch die Kamera ist ebenfalls keine Wahrnehmung der durch das LiDAR emittierten Strahlung möglich. Die Kamera misst direkt ausschließlich Intensitäten. Die Bestimmung von Entfernungsmessungen ist nur über eine Weiterverarbeitung möglich. LiDAR-Systeme messen Entfernungen und je nach Modell auch Intensitäten [37, 83, 93].

Eine Schätzung der extrinsischen Parameter zwischen einer Kamera und

einem LiDAR, die ausschließlich auf Intensitätsmessungen beruht, kann vergleichbar der Kalibrierung einer Stereokamera mittels zwei monochromer Bilder durchgeführt werden [27, 92]. Dieses Vorgehen setzt voraus, dass die Sichtfelder der verwendeten Sensoren derart umfangreich sind, dass jeweils hinreichend korrespondierende Features erkannt werden können, um eine eindeutige Zuordnung zwischen den Sensoren zu ermöglichen. Bei kleinen Sichtfeldern, z.B. eines monogonen LiDAR, führt die Beschränkung auf eine Zeile zu Mehrdeutigkeiten, welche nicht direkt aufgelöst werden können.

Ein Ansatz, welcher die Intensitätswerte der Kamera mit den Entfernungsmessungen eines monogonen LiDAR mittels eines bekannten Kalibrierobjekts verknüpft, wird von Zhang und Pless [99] vorgestellt. Für die Umsetzung wird ein bekanntes Kalibrierobjekt in beiden Datensätzen identifiziert und über den extrahierten Versatz der Datensätze die extrinsischen Parameter bestimmt. Aufbauend auf diesem Ansatz stellen Li et al. [59] eine Erweiterung vor, die gegenüber des Ansatzes von Zhang und Pless keine Punkt- sondern Kantenfeatures verwendet, was zu einer Steigerung der Genauigkeit der Schätzung führt.

3.1.2 Ermittlung der extrinsischen Parameter

Der Ansatz von Li et al. [59] geht von einem Aufbau vergleichbar zu Abbildung 3.1 aus. Ein Kalibrierobjekt, ein planares, einfarbiges, gleichschenkeliges Dreieck, wird innerhalb eines Kamerabildes und den LiDAR-Daten identifiziert. Im Kamerabild werden dazu die Eckpunkte \mathbf{a} , \mathbf{b} und \mathbf{c} lokalisiert, welche den Objektpunkten \mathbf{A} , \mathbf{B} und \mathbf{C} entsprechen. In den LiDAR-Daten werden die Schnittpunkte \mathbf{E} und \mathbf{F} bestimmt, welche auf der Schnittgeraden zwischen der Objektoberfläche des Kalibrierobjekts und der Laserebene des monogonen LiDAR die Übergänge von Objektoberfläche zu Hintergrund markieren.

Auf Basis der ermittelten Features folgen Li et al. der Annahme, dass bei einer Transformation der Objektpunkte \mathbf{E} und \mathbf{F} unter Kenntnis der ${}^c\mathbf{T}_l$ auf die Bildebene, deren Projektionen \mathbf{e} und \mathbf{f} auf den Geraden $\overline{\mathbf{ab}}$ bzw. $\overline{\mathbf{ac}}$ liegen. Folglich entspricht die Summe der orthogonalen Abstände $d(ab, e) + d(ac, f) = 0$, wobei

$$d(ab, e) = \frac{\|\vec{\mathbf{eb}} \times \vec{\mathbf{ab}}\|}{\|\vec{\mathbf{ab}}\|} \quad (3.1)$$

ist. Liegt dennoch ein Abstand vor, kann unter der Annahme, dass die intrinsische Kalibrierung der Kamera hinreichend genau ist, davon ausge-

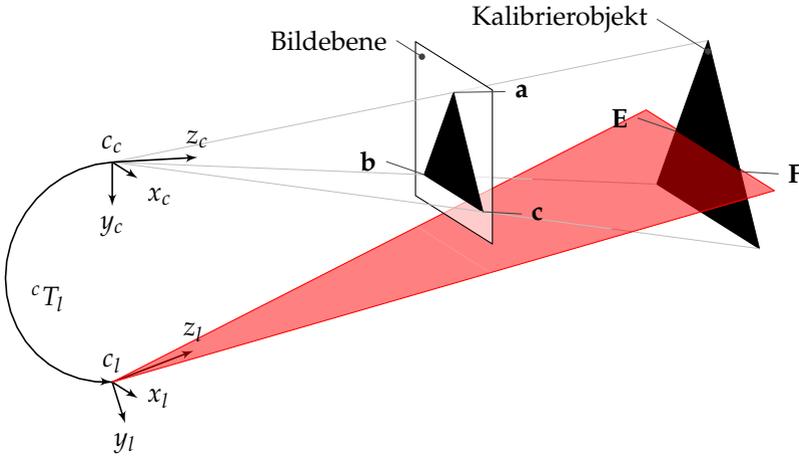


Abbildung 3.1: Aufbau einer extrinsischen Kalibrierung zwischen monokularer Kamera c und monogonem LiDAR l .

gangen werden, dass cT_l nicht korrekt ist.

Ziel der Kalibrierung ist die Ermittlung von cT_l . Gleichzeitig ist cT_l Bestandteil der Bewertung, was eine direkte Lösung verhindert. Es wird daher eine sukzessive Approximation mittels einer nichtlinearen Optimierung, etwa Methoden der kleinsten Quadrate oder Levenberg Marquardt [58, 62] genutzt, um cT_l zu ermitteln.

Das dargelegte Optimierungsproblem weist sechs Unbekannte, die Bestandteile der Orientierung $\Phi(\Theta_x, \Theta_y, \Theta_z)$ und der Translation $\mathbf{t}(t_x, t_y, t_z)$, auf. Auf der anderen Seite steht eine zu optimierende Größe, die Summe der orthogonalen Abstände. Das dargelegte Problem ist damit unterbestimmt und eine eindeutige Lösung allein auf einem Satz Kamera- und LiDAR-Daten ist nicht möglich. Für eine eindeutige Zuordnung der Parameter sind auf Grundlage der Unbekannten mindestens sechs unabhängige Datensätze erforderlich, in denen das Kalibrierobjekt in möglichst unterschiedlichen Posen betrachtet wird, um Kopplungen zwischen den einzelnen Unbekannten zu eliminieren. Die iterative Minimierung erfolgt anschließend über die Summe aller genutzten Datensätze i

$$\min \left\{ \epsilon = \sum_i \left[d(\overrightarrow{\mathbf{a}_i}, \mathbf{b}_i, \mathbf{e}_i) + d(\overrightarrow{\mathbf{a}_i}, \mathbf{c}_i, \mathbf{f}_i) \right]^2 \right\}. \quad (3.2)$$

Bei den dargelegten nichtlinearen Optimierungsverfahren handelt es sich um lokale Ansätze, so dass das Auffinden des globalen Minimums nicht

garantiert werden kann. Der Abbruch der Optimierung erfolgt über einen Schwellwert. Final ist eine manuelle Bewertung erforderlich.

Die Optimierung setzt voraus, dass die Laserebene die Sekanten des Kalibrierobjekts schneidet und bekannt ist, ob sich der Objektpunkt **A** ober- bzw. unterhalb der Laserebene befindet, da es anhand einer alleinigen Auswertung der LiDAR-Daten nicht möglich ist, eine eindeutige Zuordnung der Featurepaarungen $d(ab, e)$ und $d(ac, f)$ zu identifizieren.

3.1.3 Kantentreffer als Fehlerquelle

Der Ansatz von Li et al. [59] setzt darüber hinaus voraus, dass die vom LiDAR bereitgestellten Daten korrekt sind und die Objektpunkte **E** und **F** tatsächlich den Schnittpunkten zwischen den Sekanten und der resultierenden Schnittgeraden von Laserebene und Objektoberfläche entsprechen. Diese Annahme kann allerdings aufgrund von Kantentreffern (siehe Abschnitt 2.3.2) in Verbindung mit einer „First-Echo“-Auswertung der Entfernungsmessungen nicht garantiert werden.

Diese Einschätzung beruht unter anderem auf den Resultaten in Abbildung 3.4, für die bekannte Objekte mit unterschiedlichen Oberflächeneigenschaften und in unterschiedlichen Entfernungen mittels eines Aufbaus wie in Abbildung 3.2 vor einem identischen Hintergrund vermessen wurden. Bei einer Vermessung eines bekannten Objekts mittels des Aufbaus in

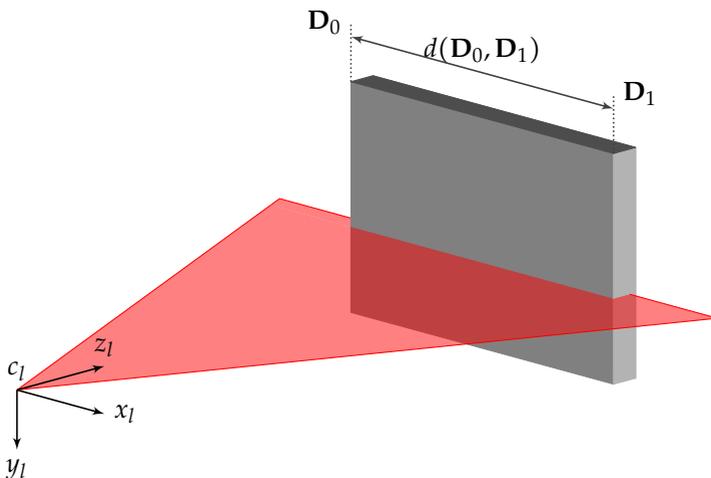


Abbildung 3.2: Vermessung einer Objektausdehnung $d(D_0, D_1)$.

Abbildung 3.2 ergeben sich Resultate für Entfernungs- und Intensitätsmessungen, wie sie exemplarisch in Abbildung 3.3 skizziert werden. Die resultierenden Entfernungsmessungen einer „First-Echo“-Auswertung sind in Abbildung 3.3 rot aufgetragen. Es ist ein planares Objekt, in ca. vier Metern Entfernung, vor einem Hintergrund in etwa sieben Metern auszumachen. Die korrespondierenden Intensitätsmessungen sind grün abgebildet.

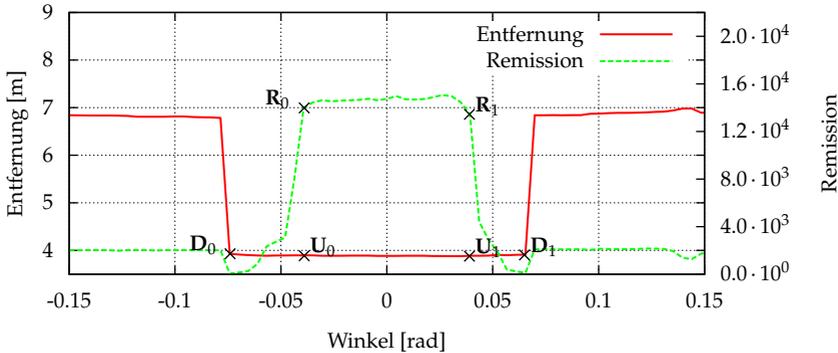


Abbildung 3.3: LiDAR Entfernungs- und Intensitätsmessungen eines bekannten Objekts [54].

Für die Identifikation der Einflüsse von Kantentreffern auf die Entfernungsmessungen wurden die Ausdehnungen $d(\mathbf{D}_0, \mathbf{D}_1)$ für drei Objekte mit unterschiedlichen Oberflächeneigenschaften ermittelt und mit manuell bestimmten Werten verglichen. Die Objektoberflächen waren planar und jeweils homogen matt weiß bzw. matt schwarz lackiert oder mit einer stark reflektierenden und bündelnden Folie beklebt, die das Licht zum überwiegenden Teil in die Richtung der Strahlenquelle reflektiert. Den resultierenden Fehler zwischen den manuell und den mittels Entfernungsmessungen ermittelten Ausdehnungen zeigt Abbildung 3.4. Die Graphen in Abbildung 3.4 zeigen, dass ein höherer Remissionsgrad (Tabelle: 2.1) zu einem größeren Fehler bei der Vermessung eines Objektes führt. Weiter ist zu erkennen, dass ein höherer Remissionsgrad zu einem stärkeren Fehlerwachstum mit zunehmender Entfernung zwischen Laser und Objekt führt. Erklären lassen sich diese Beobachtungen in Abbildung 3.4 mit der „First-Echo“-Laufzeitauswertung des verwendeten LiDAR in Verbindung mit Kantentreffern. Auf Basis der Zeitdifferenz bei Laufzeitmessungen zwischen Aussenden und Empfang wird anschließend die Entfernung bestimmt (Gleichung: 2.17). Bei einer homogenen, planaren Oberfläche ergibt sich idealerweise ein einzelnes Echo. Ist die Intensität des Echos, welches durch die

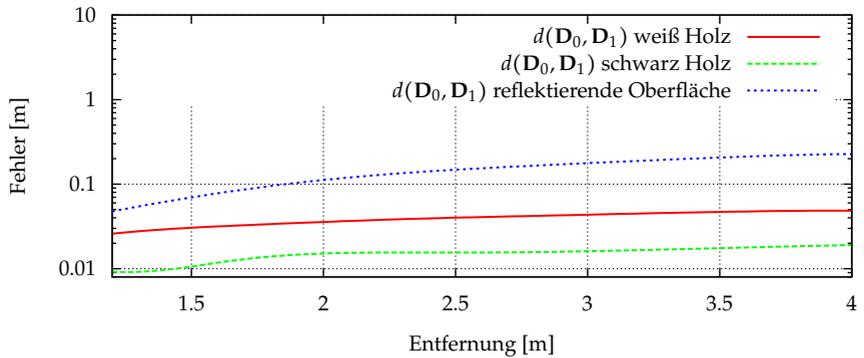
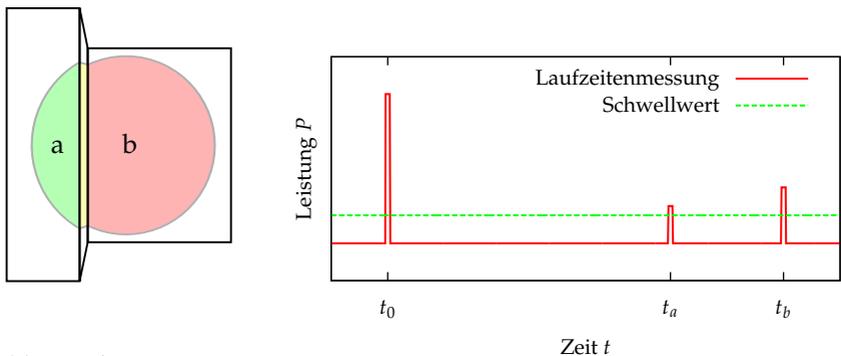


Abbildung 3.4: Fehler bei der Vermessung eines bekannten Objekts abhängig von dessen Oberflächenreflektivität, nach [54].

Entfernung, den Einfallswinkel auf der Objekt Oberfläche und den Remissionsgrad der Oberfläche beeinflusst wird, oberhalb einer Schwelle, welche die Separation zwischen Signal und Rauschen definiert, wird der Empfang eines Echos registriert.

Herrscht dagegen die Situation vor, dass der Laserstrahl eine Kante trifft



(a) Visualisierung eines Kantentreffers.

(b) Vereinfachte Darstellung der Laufzeit einer Lasermessung mit zwei Echos.

Abbildung 3.5: Kantentreffer bei „First-Echo“-Auswertung

(Abbildung: 3.5a) entstehen zwei Echos (Abbildung: 3.5b). Die Teilfläche a des Laserspots liegt auf einer Oberfläche im Vordergrund und erzeugt ein Echo zum Zeitpunkt t_a . Der Bereich b des Laserspots liegt auf einer Ober-

fläche im Hintergrund. Das Echo, das auf dem Bereich b basiert, erreicht den Detektor des Lasers zum Zeitpunkt t_b , wobei $t_b > t_a$. Für die „First-Echo“-Auswertung wird das Echo t_b aus konservativer Sicht verworfen, so dass die resultierende Entfernung ausschließlich auf t_a beruht.

Die Verwendung des Echos, welches zum Zeitpunkt t_b den Detektor erreicht, erfolgt erst an dem Punkt, an dem die Intensität des Echos t_a unter die definierte Schwelle der Rausch-Signal-Trennung fällt. Das Echo des Zeitpunktes t_b wird daher ebenfalls nicht verwendet, wenn dessen Intensität über der des Echos zum Zeitpunkt t_a liegt, wie es in Abbildung 3.5b skizziert ist.

Das Ergebnis ist, dass beim Auftreten von Kantentreffern in Kombination mit der „First-Echo“-Auswertung Objekte in den Messdaten größer erscheinen als sie in Realität sind. Dieses „Aufquellen“ ist aus konservativer Sicht im Bereich der Hindernisvermeidung durchaus erwünscht, um einen Kontakt mit dem Objekt sicher zu vermeiden. Eine genaue Vermessung eines Objektes, wie sie bei der Kalibrierung angestrebt wird, wird dadurch allerdings verhindert.

Bedingt durch den Anteil des Spots auf der Objektoberfläche ändert sich die reflektierte Lichtleistung und damit auch die flächenbedingte Intensität eines Echos. Diese Tatsache führt dazu, dass mit abnehmender Fläche, die der Laserspot auf der Objektoberfläche bestrahlt, ebenfalls die registrierte Intensität abnimmt, auch wenn die Laufzeit der Messung konstant bleibt. Die Reduktion der Intensität bei konstanten Entfernungsmessungen wird zwischen den Winkeln $\varphi_{D_0} < \varphi < \varphi_{R_0}$ und $\varphi_{R_1} < \varphi > \varphi_{D_1}$ in Abbildung 3.3 sichtbar. Die Werte R_0 und R_1 beschreiben Messungen, deren Intensitätswert knapp unterhalb des durchschnittlichen Erwartungswertes für Intensitätsmessungen im Zentrum des homogen gefärbten Objekts liegen. Die Spotflächen dieser Echos liegen noch nahezu komplett auf dem Objekt.

Eine Übertragung der Winkel φ_{R_0} und φ_{R_1} auf die Entfernungsmessung erlaubt die Identifikation der Punkte $U_0(\varphi_{R_0}, r_{D, \varphi_{R_0}})$ und $U_1(\varphi_{R_1}, r_{D, \varphi_{R_1}})$, welche eine Lokalisierung der Objektübergänge mittels Intensitätsänderungen erlauben. Eine Darstellung der Fehler zwischen dem manuell vermessenen Objekt und dem Abstand $d(U_0, U_1)$ zeigt Abbildung 3.6. Die Graphen für das weiße und schwarze Objekt dienen als Referenz und entsprechen denen aus Abbildung 3.4. Für das Objekt mit der stark reflektierenden Oberfläche wurde der Übergang zwischen Objektoberfläche und Hintergrund mittels Intensitätsmessungen bestimmt. Der Fehler zwischen erwarteter und gemessener Ausdehnung des Objekts ist für die Anwendung von $d(U_0, U_1)$ geringer als bei der Nutzung von $d(D_0, D_1)$ in Abbildung 3.4.

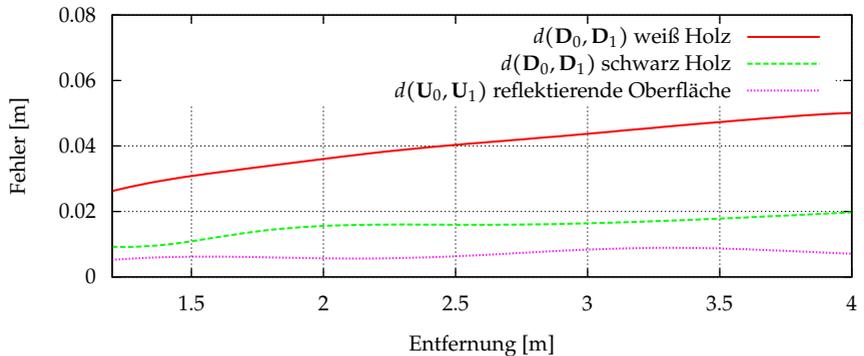


Abbildung 3.6: Fehler bei der Vermessung eines bekannten Objektes abhängig von dessen Oberflächenreflektivität, nach [54].

3.1.4 Erweitertes Kalibrierobjekt

Sowohl Zhang und Pless [99] als auch Li et al. [59] setzen auf eine manuelle Identifikation des Kalibrierobjekts in Kamera und Laserdaten, was eine zeitaufwendige und fehleranfällige Auswertung erfordert. Darüber hinaus verwenden beide ausschließlich LiDAR-Entfernungsmessungen zur Identifikation der Kalibrierobjektgrenzen, was auf Basis der Ergebnisse aus dem Abschnitt 3.1.3 nicht empfohlen werden kann.

Um einerseits die automatische Lokalisation des Kalibrierobjekts in den Kamera- und Laserdaten zu ermöglichen und andererseits eine hinreichend genaue Identifikation der Objektgrenzen mittels Laserintensitätsmessungen sicherzustellen, wird auf Basis des von Li et al. präsentierten Kalibrierobjekts ein erweiterter Aufbau vorgeschlagen.

Die Identifikation des Kalibrierobjekts mittels der Intensitätsmessungen kann, wie von Naroditsky [70] vorgeschlagen, mittels eines sprunghaften Übergangs von einer dunklen auf eine helle Oberflächenfärbung geschehen. Wobei der Intensitätssprung sowohl in den Kamerabildern als auch von Laserdaten identifizierbar sein muss.

Die von Naroditsky dargelegte Umsetzung erlaubt zwar die Detektion des Sprungs, aber keine eindeutige Zuordnung der Kante, da hell-dunkel Sprünge sehr häufig vorkommen und den Intensitäten des LiDAR keine eindeutigen Werte zugeordnet werden können. Um eine eindeutige Identifikation zu erzielen, die auch automatisch durchgeführt werden kann, ist es notwendig, eine örtliche Intensitätsdiskontinuität hervorzurufen, die derart prägnant ist, dass sie nicht zufällig oder unabsichtlich während der Kali-

brierung erzeugt werden kann.

Diese Zielstellung kann etwa durch die Nutzung eines eindeutigen Musters aus stark reflektierendem und Licht bündelndem Material erreicht werden, wobei der Remissionsgrad die Remissionsgrade der umgebenen Materialien übersteigen sollte (siehe Tabelle 2.1). Diese Materialien arbeiten unter anderem mit kleinsten Prismen, welche das eintreffende Licht gebündelt und gerichtet in Richtung der Strahlenquelle reflektieren. Das führt dazu, dass der resultierende Intensitätswert nahezu dem einer gerichteten Reflexion ohne diffuse Anteile entspricht, was für einen Strahlenverlauf innerhalb geschlossener Räume nahezu ausgeschlossen werden kann. Ausgangspunkt für das erweiterte Kalibrierobjekt bildet das homogen gefärbte, planare, gleichschenklige Dreieck, das von Li et al. verwendet wurde. Für eine gute, kontrastreiche Sichtbarkeit des schwarzen Dreiecks in den Kameradaten und um darauf aufbauend eine automatische Identifikation zu ermöglichen, wird das schwarze Dreieck entlang aller Kanten mit einem weißen Rand von circa fünf Zentimeter umrahmt. Der weiße Rahmen verhindert Irritationen zwischen dem schwarzen Dreieck und dunklen Objekten im Hintergrund, die eine erfolgreiche Identifikation in den Kameradaten behindern.

Für eine Identifikation innerhalb der Laserdaten wird das Kalibrierobjekt mit Elementen der beschriebenen Reflektorfolie ausgestattet. Auf beiden Sekanten des schwarzen Dreiecks wurde passgenau je ein zwei Millimeter starker Streifen der Reflektoroberfläche positioniert. Die Farbe der Streifen ist dunkelblau und kann aufgrund der geringen Intensitätsänderung zum schwarzen Dreieck nicht von der Kamera detektiert werden. Allein im Intensitätsbild des Lasers sind sie eindeutig sichtbar. Im Inneren des schwarzen Dreiecks befindet sich ein Schachbrettmuster, dessen Zeilen parallel zur Basiskante des Dreiecks verlaufen. Die Abbildung 3.7 zeigt eine Skizze des Kalibrierobjekts. In der Skizze sind die reflektierenden Streifen blau-weiß gestreift hervorgehoben, um zu verdeutlichen wo sie sich befinden. Durch den dargelegten Aufbau ist eine Bestimmung der Abmessungen des schwarzen Dreiecks mittels der Laserentfernungsmessungen nicht mehr möglich. Mit den Entfernungsmessungen des Lasers können ausschließlich die Abmessungen des gesamten Kalibrierobjekts ermittelt werden. Auf Grundlage der beschriebenen Nachteile (Abschnitte: 3.1.3), die bei der Bestimmung von Objektausdehnungen mittels der Entfernungsmessungen vorliegen, soll auf die Bestimmung der Objektausdehnung auf Basis der Entfernungsmessungen vollständig verzichtet werden. Die Entfernungsmessungen werden allein zum Bestimmen der Schnittgerade verwendet, die auf der Oberfläche des Kalibrierobjekts verläuft. Die Ermittlung der Schnittpunkte E und F bzw. $U_0(\varphi_{R_0}, r_{D, \varphi_{R_0}})$ und $U_1(\varphi_{R_1}, r_{D, \varphi_{R_1}})$ zwischen

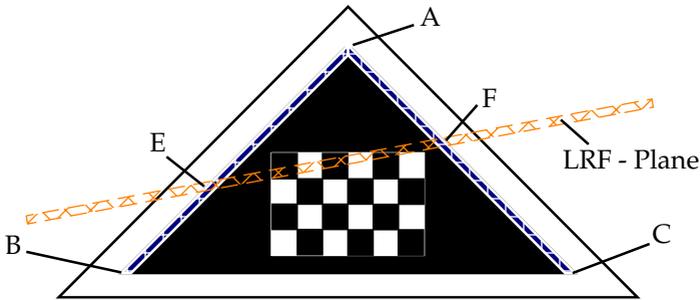


Abbildung 3.7: Skizze des Kalibrierobjekts mit gesuchten Features für Kamera und LiDAR

der Schnittgeraden und den Sekanten des Dreiecks erfolgt allein auf den Intensitätsmessungen. Dass die Sekanten des schwarzen Dreiecks nicht gleichzeitig Objektkanten sind, sondern sich innerhalb eines Objektes befinden, soll zusätzliche Irritationen bei den Entfernung- und Intensitätsmessungen durch Kantentreffer verhindern und so zu einer genaueren Bestimmung der Schnittpunkte führen.

Die Identifikation des Kalibrierobjekts bzw. der Außenkanten des schwarzen Dreiecks innerhalb des Kamerabildes erfolgt mit Kantendetektoren (siehe Abschnitt 2.4.2). Basierend auf den gefundenen Kanten bzw. deren Schnittpunkten wird das schwarze Dreieck anhand der Anzahl der Eckpunkte identifiziert. Die Schnittpunkte der gefundenen Dreieckskanten bilden die drei Ecken **a** (Spitze), **b** (linke untere Ecke) und **c** (rechte untere Ecke). Die Eckpunkte **a**, **b** und **c** auf der Bildebene entsprechen den Objektpunkten **A**, **B** und **C** im Kamerakoordinatensystem in Abbildung 3.7.

Die Identifikation des Kalibrierobjekts in den Laserdaten wird mit den Entfernung- und Intensitätsmessungen durchgeführt. Abbildung 3.8 zeigt einen beispielhaften Datensatz, in dem das Kalibrierobjekt betrachtet wird. Die zwei Spitzen R_0 und R_1 (Abbildung: 3.8) in den Intensitätsmessungen (grün) werden durch die beiden Reflektorstreifen erzeugt, die sich wie beschrieben auf den Sekanten des schwarzen Dreiecks befinden. In den Entfernungsmessungen werden die Grenzen des Kalibrierobjekts mit den abfallenden bzw. ansteigenden Flanken D_0 und D_1 zum Hintergrund sichtbar. Wie verdeutlicht, befinden sich die Reflektorstreifen nicht auf den räumlichen Kanten des Kalibrierobjekts. Die jeweilige Differenz zwischen den Winkeln φ_{D_0} und φ_{R_0} sowie φ_{D_1} und φ_{R_1} basiert daher nicht ausschließlich auf Kantentreffern, sondern lässt sich zum überwiegenden Teil auf den weißen Rand des Kalibrierobjekts zurückführen.

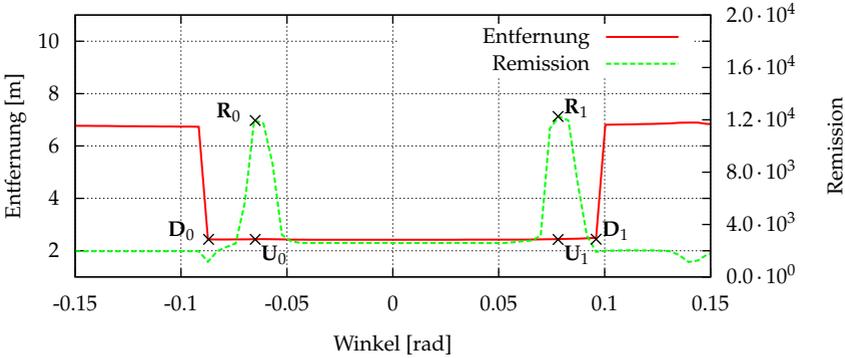


Abbildung 3.8: Entfernungs- und Intensitätsmessungen von dem Kalibrierobjekt nach Abbildung 3.7.

Ausgehend von den Winkeln φ_{R_0} und φ_{R_1} werden die Entfernungsmessungen des LiDAR, die sich zwischen diesen Winkeln befinden, zu einer Regressionsgeraden $g_{reg} : \mathbf{m} = \mathbf{m}_0 + \mathbf{n} * \lambda$ zusammengefasst. Die g_{reg} repräsentiert die Schnittgerade zwischen Laserebene und der Oberflächenebene des Kalibrierobjekts und kann bspw. mittels einer Hauptkomponenten-Analyse [51] generiert werden. Die Verwendung einer Regressionsgeraden gegenüber einer alleinigen Verwendung von Entfernungsmesspunkten hat das Ziel, den Einfluss von Rauschen und möglichen Ausreißern zu reduzieren. Die Herleitung der Strecke \overline{EF} innerhalb von g_{reg} erfolgt mittels der Schnittpunkte zwischen g_{reg} und deren zwei Normalen, die jeweils die Punkte $\mathbf{U}_0(\varphi_{R_0}, r_{D, \varphi_{R_0}})$ und $\mathbf{U}_1(\varphi_{R_1}, r_{D, \varphi_{R_1}})$ enthalten. Die Endpunkte **E** (links) und **F** (rechts) der Strecke sind in Abbildung 3.7 zu sehen. Die orthogonale Projektion wird beispielhaft mit \mathbf{U}_0 und dem resultierenden Punkt

$$E = \mathbf{m}_0 + \mathbf{n} * \left(\frac{\mathbf{n} * (\mathbf{u}_0 - \mathbf{m}_0)}{\mathbf{n}^2} \right) \quad (3.3)$$

gezeigt. Die verwendeten Koordinaten \vec{u}_0 liegen kartesisch in \mathcal{R}^3 vor und entsprechen \mathbf{U}_0 in Polarkoordinaten. Die Punkte **E** und **F** bzw. deren Projektion bilden mit den identifizierten Eckpunkten **a**, **b** und **c** des Dreiecks den Ausgangspunkt für die Kalibrierung nach Li et al. in Abschnitt 3.1.2.

3.1.5 Validierung

Für die Validierung des hier vorgestellten Ansatzes zur Bestimmung der extrinsischen Parameter zwischen monokularer Kamera und monogonem LiDAR werden zwei Aspekte betrachtet. Auf der einen Seite steht die Fähigkeit des Ansatzes, das Kalibrierobjekt in den Kamera- und Laserdaten automatisch zu detektieren, im Fokus. Auf der anderen Seite wird die Genauigkeit bei der Bestimmung der extrinsischen Parameter bewertet und in Relation zu den bereits bestehenden Ansätzen von Li et al. [59] und Zhang et al. [99] gestellt.

Automatische Detektion des Kalibrierobjekts

Die Fähigkeit, das Kalibrierobjekt in den Laser- und Kameradaten automatisch zu detektieren, wird beispielhaft anhand eines Büroszenarios gezeigt.

Versuchsaufbau Für die Versuchsdurchführung wurde eine beliebige Büroumgebung innerhalb eines Gebäudes mit künstlichen (Lampen, Monitore) und natürlichen (Fenster) Lichtquellen gewählt. Bei der Wahl der Kameraeinstellungen für Blende und Verschlusszeiten wurde darauf geachtet, dass eine gute Sichtbarkeit der betrachteten Umgebung sowie eine hinreichende Tiefenschärfe vorliegt. Die Wahl der Einstellungen erfolgte manuell.

Der LiDAR verwendet eine „First-Echo“-Auswertung für die Entfernungsmessungen und stellt pro Entfernungsmessung eine Intensität bereit. Es wurden keine Anpassungen an dem LiDAR vorgenommen.

Beide Sensoren sind statisch mit einander verschraubt und so zu einander ausgerichtet, dass sich Teile ihres jeweiligen Sichtfeldes überschneiden. Die Pose der Sensoren zueinander ist nur bezüglich der Sensorgehäuse bekannt. Dies ist, wie in Abschnitt 3.1 ausgeführt eine Fehlerquelle, da die Position des Sensors im Gehäuse unbekannt ist.

Das Kalibrierobjekt ist auf einem Stativ montiert. Das Kalibrierobjekt ist identisch zu dem in Abschnitt 3.1.4.

Die Aufnahmen der einzelnen Sensoren werden über Zeitstempel synchronisiert. Die Synchronisierung erfolgt über die Zeit des Rechners, an den beide Sensoren angeschlossen sind. Auf diesem Rechner erfolgt auch die Berechnung der automatischen Detektion und die Bestimmung der extrinsischen Parameter.

Versuchsdurchführung Zum Testen der automatischen Detektion wurden von dem Kalibrierobjekt, pro Versuch, 30 korrespondierende Aufnah-

men aus unterschiedlichen Posen erstellt. Die Identifikation des Kalibrierobjekts erfolgte unverzüglich mittels des vorgestellten Ansatzes. Zum Aufbau einer umfangreicheren Datenbasis wurden 20 Versuche durchgeführt [54].

Versuchsergebnisse Abbildung 3.9 zeigt eine beispielhafte Aufzeichnung für Intensitäts- und Entfernungsmessungen des LiDAR und das korrespondierende Kamerabild. Die Ebene in der die Messungen des LiDAR liegen ist in diesem Beispiel etwa parallel zum abgebildeten Boden des Raums.

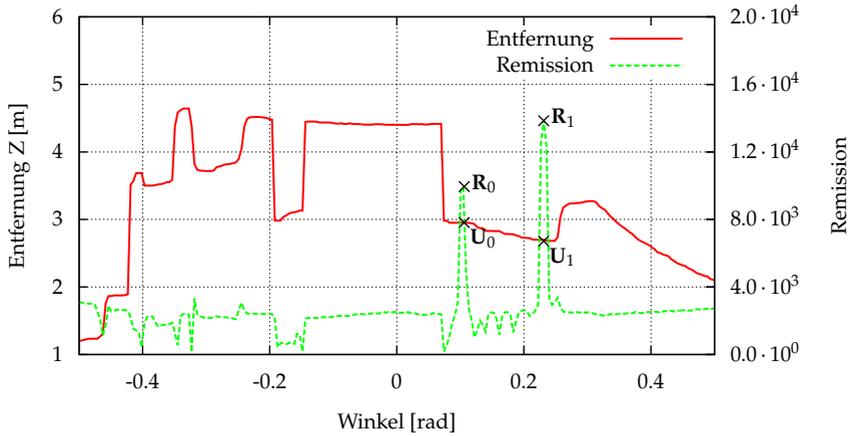
Die Entfernungsmessungen des LiDAR zeigen somit einen Schnitt durch den Raum, der von links nach rechts die zwei Wasserflaschen, die hintere Stuhllehne, die vordere Stuhllehne, die Wand im Hintergrund mit Tür, das Kalibrierobjekt und den Schrank schneidet (siehe zum Vergleich das Kamerabild).

Die Intensitätsmessungen zeigen jeweils die korrespondierende Remission pro Entfernungsmessung. Auffällig sind zwei Spitzen $R_0 \approx 10000$ und $R_1 \approx 14000$, welche die Schnittpunkte der Laserebene mit den Reflektorstreifen auf dem Kalibrierobjekt repräsentieren. Die Punkte U_0 und U_1 bilden die errechneten Schnittpunkte bzw. Entfernungsmessungen der Laserebene mit den Sekanten des schwarzen Dreiecks innerhalb des Kalibrierobjekts.

In dem Kamerabild zeigen die Markierungen die identifizierten Ecken des Kalibrierobjekts.

Ergebnisse zu der Genauigkeit der bestimmten extrinsischen Parameter dieses Kalibrierversuchs werden im folgenden Abschnitt „Genauigkeit der extrinsischen Parameter“ unter der Bezeichnung „Experiment“ vorgestellt.

Diskussion Die automatische Identifikation der Messungen U_0 und U_1 in den LiDAR-Daten sowie der Dreiecksecken a , b und c im Kamerabild, stellen die benötigten Features bereit (vergleiche Abschnitt 3.1.2), um die automatische Ermittlung der extrinsischen Parameter zu ermöglichen. Unter Vernachlässigung der Intensitätsmessungen ist eine Identifikation des Kalibrierobjekts (U_0, U_1) in den Entfernungsmessungen nur über eine manuelle Plausibilitätsprüfung mit dem Kamerabild möglich. Eine eindeutige Bestimmung des Kalibrierobjekts allein auf der Entfernungsmessung ist dagegen schwierig. Beim Einbeziehen der Intensitätsmessungen R_0 und R_1 , welche eine beispielhafte Grenze von ≈ 4000 auf der Skala der Intensitätsmessung deutlich übersteigen, ist eine eindeutige Identifikation des Kalibrierobjekts auch in komplexeren Umgebungen möglich. Die Automatisierung dieses Arbeitsschrittes ist die Grundlage zur Bearbeitung von



(a) Entfernungs- und Intensitätsmessungen des hier verwendeten LiDAR.



(b) Monokulares, monochromes Kamerabild.

Abbildung 3.9: Kalibrierobjekt in einem Büro [54].

größeren Mengen an Lasermessungen und ist die Basis für die Kombination mehrerer Lasermessungen mit einem Kamerabild. Diese Kombination ist ein Baustein der präsentierten extrinsischen Parameterbestimmung und führt zu einer Steigerung der Kalibrierengenauigkeit, was in Abbildung 3.10 gezeigt wird.

Genauigkeit der extrinsischen Parameter

Zur Bewertung der Genauigkeit, mit welcher die extrinsischen Parameter in dem hier vorgestellten Ansatz bestimmt werden, werden die generierten Ergebnisse mit den Resultaten bereits bestehender Arbeiten von Zhang [99] und Li et al. [59] verglichen. Die Gegenüberstellung erfolgt mit Simulationsergebnissen sowie den bestimmenden Parametern des realen Aufbaus aus dem Abschnitt „Automatische Detektion des Kalibrierobjekts“.

Versuchsaufbau In der Simulation werden mögliche Aberrationen der Sensoren nicht berücksichtigt. Die Simulationsergebnisse stellen damit das Optimum dar, welches unter Berücksichtigung der gegebenen Sensoren theoretisch erreicht werden kann.

Die in der Simulation genierten optimalen Rohdaten der Sensoren werden additiv mit einem Rauschen versehen, um die Realität besser abzubilden. Um einerseits eine Vergleichbarkeit zwischen den simulativen und experimentellen Ergebnissen zu ermöglichen, wurde ein Simulationsdurchlauf mit einer intrinsischen Kameramatrix k_{VGA} umgesetzt, welche auf die reale Kamera in Tabelle 4.1 zurückzuführen ist. Auf der anderen Seite wird eine Vergleichbarkeit zu den Resultaten von Li et al. [59] angestrebt, indem die dort beschriebene XGA Kamera ($1024px \times 768px$) mit den verwendeten Rauschparametern in der Simulation umgesetzt wurde.

Die Kameraunschärfe wurde mit einer Gauß-Verteilung abgebildet, deren Erwartungswert null und deren Standardabweichung $\sigma = 0.5px$ entspricht. Die Unschärfe wird bildseitig auf die Projektion in der simulierten Kamera additiv angewendet. Die verwendeten Parameter für die Kameraunschärfe sind für alle Simulationsdurchläufe identisch.

Die Simulation des LiDAR basiert auf keinem physikalischen Modell. Es werden einfache mathematische Strahlen verwendet. Die Entfernungsmessungen sind somit nicht mit Aberrationen, wie Kantentreffern oder dem Einfluss des Eintrittswinkels auf der Objektfläche, kontaminiert. Des Weiteren enthält die Simulationsumgebung nur ein Kalibrierobjekt, was eine automatische Detektion überflüssig macht, da die Identifikation immer eindeutig ist. Da weder eine Identifikation des Kalibrierobjekts notwendig ist, noch die Entfernungsmessungen verfälscht sind, liegen in der

Simulation keine Intensitätsmessungen für den Laser vor. Die optimalen Entfernungsmessungen in der Simulation sind mit einem gleichverteilten Rauschen kombiniert, dessen Erwartungswert null und dessen Standardabweichung $\sigma = 0.05 \text{ m}$ ist. Die gewählten Parameter des Rauschens entsprechen denen des in Tabelle 4.1 beschriebenen realen Sensors und sind identisch mit den Annahmen in den Arbeiten von Li et al. und Zhang et al..

Das Kalibrierobjekt in der Simulation ist ein einfaches gleichschenkliges monochromes Dreieck, welches im Vergleich zum realen Kalibrierobjekt keine weiteren Merkmale aufweist. Es entspricht daher dem Kalibrierobjekt, welches Li et al. [59] verwendeten. Die Position und die Orientierung des Kalibrierobjekts sind beliebig. Die einzige Beschränkung ist, dass es von beiden Sensoren simultan vollständig erfasst werden kann.

Der Sensoraufbau von Kamera und LiDAR ist statisch und die Ausgangsparameter für die extrinsische Kalibrierung sind $\Phi(0^\circ, 0^\circ, 0^\circ)$ und $\mathbf{t}(0 \text{ m}, 0 \text{ m}, 0 \text{ m})$.

Versuchsdurchführung In der Simulation wurden Stichprobengrößen von n =sechs bis $n = 30$ korrespondierenden Aufnahmen des Kalibrierobjekts herangezogen, um die extrinsischen Parameter einer Sensorkonfiguration zu bestimmen. Für jede Stichprobengröße wurden 500 Testläufe automatisiert durchgeführt und die Ergebnisse gemittelt. Welche Posen das Kalibrierobjekt einnimmt ist ebenso wie das Rauschen der beiden Sensoren pro Versuch zufällig.

Zur Validierung des vorgestellten Ansatzes wurden drei Simulationen und ein reales Experiment durchgeführt. Die Simulationsergebnisse von Zhang [99] wurden mit dem Versuch „sim. VGA Kamera“ ebenso wie die Ergebnisse von Li et al. [59] mit dem Versuch „sim. XGA Kamera“ reproduziert. Der Versuch „sim. VGA mit mean LiDAR & Subpixel“ zeigt die Simulationsergebnisse des hier vorgestellten Ansatzes zur automatisierten Bestimmung der extrinsischen Parameter. Die Bezeichnung „mean LiDAR“ verdeutlicht die Erstellung einer Regressionsgeraden mittels mehrerer Lasermessungen, welche auf der Schnittkante zwischen Objektfläche und Laserebene liegen. Die Schnittpunkte der Laserebene mit den Sekanten des Kalibrierobjekts werden anschließend auf diese Gerade projiziert. Diese Umsetzung folgt dem vorgestellten Ansatz, bei dem die Schnittpunkte $\mathbf{U}_0(\varphi_{\mathbf{R}_0}, r_{\mathbf{D}, \varphi_{\mathbf{R}_0}})$ und $\mathbf{U}_1(\varphi_{\mathbf{R}_1}, r_{\mathbf{D}, \varphi_{\mathbf{R}_1}})$ auf eine Regressionsgerade projiziert werden, welche mittels Entfernungsmessungen erstellt wurde (Abschnitt 3.1.4). Die verwendeten Lasermessungen der Regressionsgeraden wurden mit dem beschriebenen Rauschen versehen. Die automatisierte Detektion des Kalibrierobjekts im präsentierten Ansatz ermöglicht es, in den Kameradaten das Kalibrierobjekt mit einer Subpixel-Genauigkeit von $1/10 \text{ px}$ zu

bestimmen. Diese Fähigkeit war bei den Umsetzungen von Li et al. und Zhang [59, 99] nicht gegeben und wird ebenfalls in die Simulation überführt.

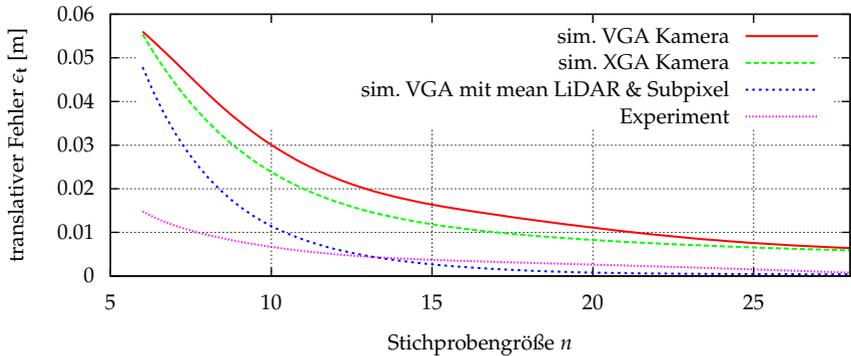
Das Experiment verdeutlicht die Resultate der Kalibrierung, welche mit realen Sensoren und dem realen Kalibrierobjekt erreicht wurden und repräsentiert die reale Umsetzung der Simulation „sim. VGA mit mean LiDAR & Subpixel“.

Für die Simulation ist die Ground Truth, die genaue Transformation zwischen der Kamera und Laser, bekannt. Dadurch ist eine hinreichend genaue Fehlerabschätzung möglich. Für die realen Versuche ist die Ground Truth nur bezüglich der Sensorgehäuse bekannt. Diese ist, wie in Abschnitt 3.1 ausgeführt, eine Fehlerquelle, da die Position des Sensors im Gehäuse unbekannt ist. Eine Methode zur Bestimmung der „wahren“ Sensorposition lag zum Zeitpunkt der Experimente nicht vor. Aus diesem Grund wurde, nach einer Prüfung auf Plausibilität, für die Messungen der 20 durchgeführten realen Testläufe, welche mit einem identischen Aufbau durchgeführt wurden, der Median ermittelt und als Ground Truth verwendet. Welche Posen das Kalibrierobjekt während des Kalibriervorgangs gegenüber dem Sensoraufbau einnimmt ist ebenfalls beliebig.

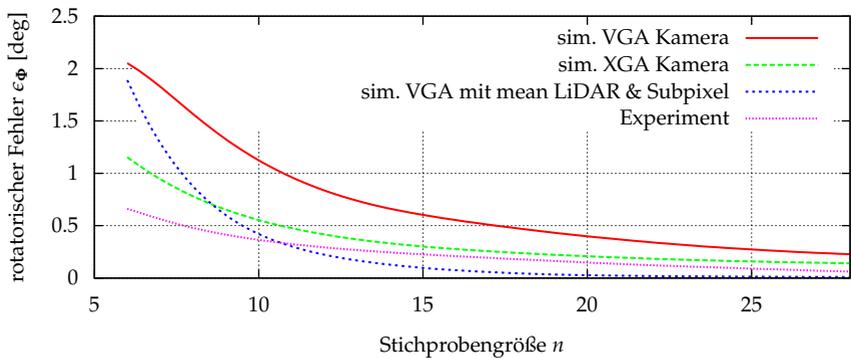
Versuchsergebnisse Die Ergebnisse der Simulation und der realen Kalibrierung zeigt die Abbildung 3.10. Die skalaren Fehlerwerte ϵ_t und ϵ_ϕ pro Graph in Abbildung 3.10 ergeben sich wie folgt: Jede einzelne Kalibrierung generiert zwei dreidimensionale Vektoren, welche die Abweichung der Position ϵ_t und der Lage ϵ_ϕ zwischen den approximierten extrinsischen Parametern und der Ground Truth beschreiben. Für jede der generierten Größen, Position- oder Lagefehler wird über die jeweilige Anzahl an Testläufen der Mittelwert gebildet. Der resultierende Vektor wird schlussendlich mittels Mittelwertbildung über die Bestandteile des Vektors zu einem Skalar zusammengefasst.

Der Graph „sim. VGA Kamera“ verdeutlicht die Resultate, welche mit einer VGA Kamera erreicht werden können, wenn keine weiteren Optimierungsoptionen genutzt werden. Die gewählte Umsetzung entspricht in ihren Parametern denen von Zhang [99]. Der Graph „sim. XGA Kamera“ zeigt die Resultate einer simulierten Kalibrierung mit einer XGA Kamera ebenfalls ohne den Einsatz jeglicher Optimierung. Die gewählten Parameter entsprechen denen der Arbeit von Li et al. [59].

Der Graph „sim. VGA mit mean LiDAR & Subpixel“ zeigt die Simulationsergebnisse des hier vorgestellten Ansatzes zur automatisierten Bestimmung der extrinsischen Parameter. Der Graph „sim. VGA mit mean LiDAR & Subpixel“ weist ab einer Stichprobengröße von $n \approx 8$ korrespon-



(a) Gemittelter translative Fehler, nach Krause [54].



(b) Gemittelter rotatorische Fehler, nach Krause [54].

Abbildung 3.10: Skalarer Fehler der Kalibrierung gegenüber der Stichprobengröße.

dierenden Sensordatenpaaren einen geringeren translatorischen ϵ_t als auch rotatorischen ϵ_Φ Fehler als die reproduzierten Ergebnisse der Arbeiten von Zhang und Li et al. auf.

Der Graph „Experiment“ verdeutlicht die Resultate der Kalibrierung, welche mit realen Sensoren und dem realen Kalibrierobjekt erreicht wurden und repräsentiert die reale Umsetzung des Graphen „sim. VGA mit mean LiDAR & Subpixel“.

Diskussion Eine quantitative Gegenüberstellung der Graphen „sim. VGA Kamera“ und „sim. XGA Kamera“ mit den jeweiligen Arbeiten [59, 99] zeigt eine gute Übereinstimmung der Resultate. Alle Graphen zeigen mit zunehmender Größe der Stichprobe eine Reduktion der Kalibrierfehler ϵ_t und ϵ_Φ . Sowohl die Umsetzungen von Li et al. und Zhang als auch die hier vorgestellten Reproduktionen ihrer Ansätze erreichen bei einer Stichprobengröße von $n = 25$ korrespondierende Aufnahmen $\epsilon_t \approx 0.005 m$.

Bei den rotatorischen Fehlern können die Resultate der vorangehenden Arbeiten nicht exakt reproduziert werden. Die Reproduktionen erreichen bei $n = 25$, $\epsilon_\Phi \approx 0.27^\circ$ für die VGA Kamera, was etwa einer Halbierung des Fehlers gegenüber Zhang entspricht und $\epsilon_\Phi \approx 0.15^\circ$ für die XGA Kamera, was etwas oberhalb des von Li et al. ermittelten Fehlers liegt. Die Reduktion von ϵ_Φ bei der VGA Kamera ist wahrscheinlich auf die Verwendung von Kantenfeatures anstelle Punktfeatures zurückzuführen, was dem Ziel von Li et al. entsprach. Für die Steigerung von ϵ_Φ bei der XGA Kamera gab es zum Zeitpunkt der Auswertung keine Deutung.

In der Summe werden die Reproduktionen der Ansätze von Li et al. und Zhang [59, 99] als erfolgreich betrachtet und es wird davon ausgegangen, dass diese Ergebnisse den Stand der Technik vor dem hier vorgestellten Ansatz darstellen.

Die Graphen der Simulation „sim. VGA mit mean LiDAR & Subpixel“ und des realen Testaufbaus „Experiment“ zeigen kleinere Abweichungen gegenüber ihren jeweiligen Ground Truths als die Ergebnisse der bisherigen Ansätze, was für eine Steigerung der Genauigkeit bei dem hier vorgestellten Ansatz gegenüber dem bisherigen Stand der Technik spricht. Die Abweichungen des Experiments weisen bei größeren Stichprobengrößen einen größeren Fehler als die Simulation auf, was wahrscheinlich auf die Abwesenheit von Aberrationen in der Simulation zurückzuführen ist.

Auffällig in Abbildung 3.10 ist, dass die Ergebnisse vom Graph „Experiment“ bei $n = 6$ geringere Fehler aufweisen als die Ergebnisse der zuvor besprochenen Simulationsläufe. Ursache könnte sein, dass die Transformation für die realen Sensoren recht nahe bei der inertialen Lösung von $\Phi = [0^\circ, 0^\circ, 0^\circ]^T$ lag. Allein die Translation war um wenige Zentimeter, klei-

ner zehn, verschoben. Die Simulationen wurden mit verschiedensten Ausgangskonfigurationen durchgeführt, wobei die reale Konfiguration ausschließlich mit einem Aufbau umgesetzt wurde.

Die experimentellen Ergebnisse zeigen tendenziell das gleiche Resultat wie die Simulationen. Mit zunehmender Stichprobengröße reduziert sich die Abweichung zu den Erwartungswerten. Die Ursachen für den geringen Fehler des hier vorgestellten Ansatz gegenüber den Arbeiten von Zhang [99] und Li et al. [59] können einerseits die Verwendung einer Laserschnittgeraden gegenüber dem Schnittpunkt bei Li et al. und Zhang oder andererseits die höhere Kameraauflösung durch die Verwendung von Subpixeln sein. Die Gegenüberstellung der Graphen „sim. VGA Kamera“ und „sim. XGA Kamera“ lässt vermuten, dass eine Steigerung der Kameraauflösung nur einen geringen bis keinen Einfluss auf den translatorischen Fehler ϵ_t hat, wobei durchaus eine Reduktion des rotatorischen Fehlers ϵ_ϕ zu beobachten ist. Überträgt man diese Beobachtung auf die Resultate von „sim. VGA mit mean LiDAR & Subpixel“, sollte die Verwendung von Subpixel-Genauigkeit auf ϵ_t ausschließlich einen marginalen Einfluss haben, da es sich dabei grundsätzlich um eine Steigerung der Auflösung handelt. Dennoch ist eine Reduktion von ϵ_t zu beobachten. Eine Erklärung für die Reduktion könnte sein, dass die gesteigerte Genauigkeit bzw. die Reduktion von Rauschen bei den Lasermessungen, welche durch die Kombination mehrerer Messungen zu einer Schnittgeraden entsteht, einen größeren Einfluss auf eine Reduktion von ϵ_t hat.

Fazit

In den Versuchen konnte simulativ gezeigt werden, dass die Resultate des hier präsentierten Ansatzes die Resultate von zuvor durchgeführten Arbeiten sowohl im Bereich der translatorischen als auch rotatorischen Fehler unterschreiten. Eine experimentelle Bewertung war nur eingeschränkt möglich. Die Ergebnisse zeigen qualitativ aber ein vergleichbares Verhalten zur Simulation. Darüber hinaus konnte in den Experimenten gezeigt werden, dass die automatische Detektion des Kalibrierobjekts mittels der Intensitätsmessungen des LiDAR sehr gut funktioniert. Als Folge können automatisch mehrere Lasermessungen mit einem Kamerabild kombiniert werden, was wiederum zu einer Steigerung der Kalibrierergebnisse führt.

3.2 Diskretisierung der Umwelt

Der vorgestellte Ansatz Sichtfelderweiterung arbeitet sowie zahlreiche andere Hindernisdetektionen [7, 46, 79] mit einer globalen, metrisch diskreti-

sierten Hindernisrepräsentation, im Speziellen einem Belegtheitsgitter (Abschnitt 2.5). Eine Zielstellung für den Einsatz eines Belegtheitsgitter ist die Fusion von einzelnen Sensormessungen, um einerseits einen umfassenden Blick auf die Umwelt zu erhalten, welcher über das Sichtfeld eines einzelnen Sensor hinaus geht, und andererseits, um Unzulänglichkeiten der Sensoren wie Messrauschen oder gar Fehlmessungen durch ein probabilistisches Modell in ihrer Aussagekraft zu reduzieren.

Bei der Übertragung der kontinuierlichen Umwelt in die diskrete Repräsentation M werden Raumdiskrete/Zellen m auf Basis der Sensormessungen durch Hindernisse belegt bzw. als frei deklariert und können je nach Entscheidung durchquert oder sollten gemieden werden. Aufgrund der Auflösung des Belegtheitsgitter M bzw. der Abtastauflösung des Sensors kann es bei der Übertragung in M zu widersprüchlichen Aussagen kommen, da für eine Zelle m mit einer definierten Größe nicht garantiert werden kann, dass das gesamte Volumen, welches m beschreibt, in der Realität vollständig von einem Objekt belegt bzw. frei ist. In einer Situation, in der mehrere Messungen eines Sensors die Zelle m teilweise belegen bzw. schneiden, entstehen somit konkurrierende Belegtheitswahrscheinlichkeiten $p(Z, Z_{p_n})$ für m .

Dass auf Basis gegensätzlicher Belegtheitswahrscheinlichkeiten Belegungen von Zellen revidiert werden, ist durchaus ein angestrebtes Ziel eines Belegtheitsgitters, um etwa dynamische Objekte in Repräsentation zu berücksichtigen. Allerdings haben die konkurrierenden Belegtheitswahrscheinlichkeiten auch Einfluss auf die Inklusion von statischen Hindernissen. Das Resultat ist, dass bedingt durch:

- die Auflösung des Gitters
- Winkel und Entfernungsauflösung der Entfernungsmessungen
- den Verlauf der Entfernungsmessungen bzgl. der Gitterachsen

teilweise mit Objekten belegte Zellen freigegeben werden. Die Folge könnten Kollisionen des UA mit „übersehenen“ Objekten sein.

Zur Einführung in die Thematik wird auf das Problem der bedingten Freizeichnung bei statischen Objekten, wie sie bei der Inklusion einer kontinuierlichen Welt in ein diskretes Gitter entstehen kann, eingegangen. Darauf aufbauend werden drei Ansätze vorgestellt, wie das bedingte Freizeichen von teilweise belegten Zellen reduziert werden kann. Schlussendlich erfolgt eine Parameteridentifikation pro Ansatz sowie eine Gegenüberstellung der drei Ansätze.

3.2.1 Problemdarstellung

Zur Verdeutlichung der ungewollten Freizeichnung bei teilweise belegten Gitterzellen wird von einem vereinfachten, zweidimensionalen Aufbau, in Abbildung 3.11a ausgegangen. Von einer konstanten Sensorposition O_s werden drei Entfernungsmessungen \mathbf{p}_0 - \mathbf{p}_2 hin zu einer Oberfläche (blau) erzeugt. Die Messungen werden anschließend in ein Belegtheitsgitter M mit einer Zellgröße von d_m überführt. Aufbauend auf der Situation in Ab-

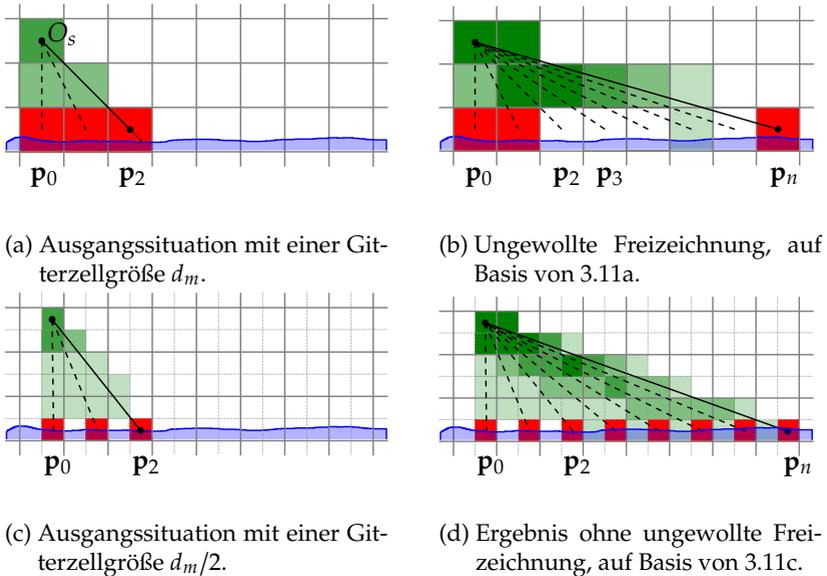


Abbildung 3.11: Freizeichnung von belegten Zellen durch zu grobe Auflösung des Belegtheitsgitters.

Abbildung 3.11a werden weitere Messungen \mathbf{p}_3 - \mathbf{p}_n in das Gitter inkrementell integriert (siehe Abschnitt 2.5.2). Die Integration der Messungen führt zu einer inkrementellen Aktualisierung der Belegungswahrscheinlichkeit in betroffenen Zellen. Das Ergebnis zeigt Abbildung 3.11b, wobei freie Zellen grün, belegte Zellen rot und unbekannte Zellen weiß markiert sind.

Am Beispiel der Messungen \mathbf{p}_2 und \mathbf{p}_3 wird die Diskrepanz von konkurrierenden Messungen für eine Zelle sichtbar. Abbildung 3.11a zeigt die Integration der Messung \mathbf{p}_2 in das Gitter, wobei sich die Zielkoordinaten in der Zelle $m_{\mathbf{p}_2}$ befinden und als „belegt“ gekennzeichnet wird. In Abbildung 3.11b wird die Messungen \mathbf{p}_3 in das Gitter integriert. Die Koordinaten von \mathbf{p}_3 liegen nicht innerhalb $m_{\mathbf{p}_2}$. Der Strahl, der \mathbf{p}_3 mit O_s verbindet,

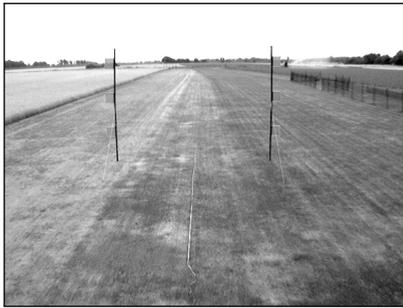
schneidet allerdings m_{p_2} . Dass p_2 in m_{p_2} endet und p_3 diese Zelle schneidet, führt zu einer widersprüchlichen Aussage über die Belegung von m_{p_2} . Bei einer gleichen Gewichtung der Aussagen „belegt“ und „frei“ ist keine eindeutige Entscheidung über die Belegung der Zelle möglich, was zu dem Resultat „unbekannt“ führt. Je nach Anzahl der Messungen können ursprünglich als „belegt“ gekennzeichnete Zellen als „frei“ deklariert werden obwohl sie teilweise von einem Objekt belegt sind. In Abbildung 3.11b wird dieses Problem von den Zellen m_{p_2} bis ausschließlich m_{p_n} verdeutlicht.

Das gezeigte Phänomen ist auf den Alias-Effekt [8, Kapitel 7.12] zurückzuführen und basiert auf dem Grad der Diskretisierung einer analogen Umwelt durch ein Gitter.

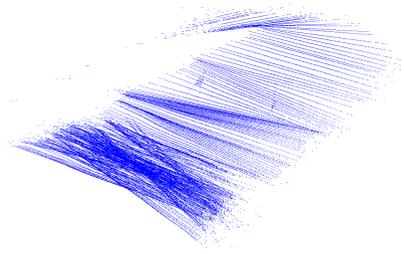
Abbildung 3.11c und 3.11d zeigen den identischen Aufbau zu Abbildung 3.11a bzw. 3.11b, allerdings mit einer Zellenkantenlänge von $d_m/2$. Die Steigerung der Gitterauflösung ist vergleichbar mit einer Steigerung der Abtastfrequenz. Das führt dazu, dass in diesem Beispiel, unter Berücksichtigung der verwendeten Winkelauflösung des Sensors und der jeweiligen Messentfernung, eine als belegt gekennzeichnete Zelle jeweils ausschließlich von einem Messstrahl geschnitten wird und somit keine Freizeichnung von teilweise belegten Zellen auftritt. Eine Steigerung der Gitterauflösung sollte allerdings nicht unbegrenzt durchgeführt werden, da jede Halbierung der Zellenkantenlänge zu einer kubischen Steigerung der Zellenanzahl führt. Darüber hinaus löst eine Steigerung der OcTree-Auflösung das Problem nur bedingt. Durch eine Änderung des Einfallswinkels der Strahlen auf die approximierte Objektfläche oder eine höhere Winkelauflösung des Sensors können wiederum ungewollte Freizeichnungen auftreten.

Die in dem Beispiel 3.11 verwendete Umsetzung des inkrementellen Updates einer Gitterzelle ist vereinfacht und dient zur Verdeutlichung des Sachverhaltes. Es wird von einer gleichen Gewichtung bei der Belegung und Freizeichnung einer Zelle ausgegangen. Bei einer konkreten Umsetzung mittels des Bayes-Updates 2.60 sind im Allgemeinen mehrere Strahlen notwendig, um eine belegte Zelle wieder als frei zu deklarieren.

Ein Beispielszenario, bei dem eine Wiese mit zwei Hindernissen von einem LiDAR vermessen wird, welches von einem unbemannten Hubschrauber getragen wird, zeigt Abbildung 3.12. Die Auftrittswinkel der Laserstrahlen auf die Ebene der Wiese sind eher groß ($\approx 70^\circ$), was dazu führt, dass die Zellen mit $d_m = 0,5m$ mehrfach von Strahlen geschnitten und aktualisiert werden. Als Folge sind die Resultate einiger Zellen widersprüchlich, was zu einer Deklaration von Freiraum für einige, den Boden beschreibenden, Zellen führt. Die Repräsentation des Bodens durch den OcTree weist



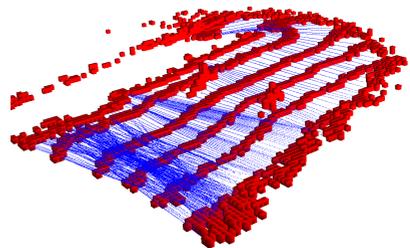
(a) Kamerasicht.



(b) Punktwolke, generiert mit einem monogonen Laser.



(c) Erwartetes Belegtheitsgitter.



(d) Tatsächlich generiertes Belegtheitsgitter.

Abbildung 3.12: Beispielszenario für ungewollte Freizeichnung.

folglich Bereiche auf, die bei einer alleinigen Stützung auf den OcTree als passierbar deklariert werden, was bei Unterlassung einer Verifikation mit der Realität zu einer Kollision des Sensorträgers mit dem Boden führen würde.

Für eine vollständige Lösung des Alias-Effekts sollte die Gitterauflösung, angelehnt an das Nyquist-Shannon-Abtasttheorem, an der Koordinate, an welcher der Strahl das Objekt trifft, mindestens doppelt so hoch sein wie die Sensorauflösung. Um eine Berechenbarkeit, etwa in Echtzeit-Anwendungen, sicherzustellen, ist die Verwendung einer derartigen umfangreichen OcTree-Auflösung allerdings nicht in allen Umsetzungen möglich. Aus diesem Grund wurden Umsetzungen entwickelt, welche die Auswirkungen den Alias-Effekt ebenfalls bei Gitterauflösungen, die unterhalb des Auflösung des Sensors liegen, reduzieren. Das bedeutet, dass der Alias-Effekt zwar immer noch vorhanden ist, allerdings weniger sichtbar ist als in Abbildung 3.12d.

3.2.2 Reduktion ungewollter Freizeichnung mittels großer Punktmengen

Der beschriebene Sachverhalt wird ebenfalls von Scherer [79] und Wurm [98] dargelegt. Als Lösungsansatz wird vorgeschlagen, dass nicht einzelne Messungen, sondern eine größere Menge von Messungen, etwa ein kompletter Scan oder eine Punktwolke C_i , im Vorfeld kombiniert und anschließend dem OcTree oder Gitter hinzugefügt werden. Über die Anzahl der Messungen \mathbf{p} innerhalb C_i oder den Umfang der Überlappung zwischen den einzelnen C_i , werden allerdings keine Aussagen getroffen.

Dieses Vorgehen wird ebenfalls von Andert [7, S. 44] umgesetzt, dessen Ansatz im Folgenden erläutert wird. Basierend auf der Datenmenge eines Stereotiefenbildes ($640\text{px} \times 480\text{px}$ [7, S. 96]) zum Zeitpunkt t wird ein separates Belegtheitsgitter M_t generiert. Bei der Umsetzung des Objektpunkts \mathbf{p} in M_t wird allerdings kein Bayers-Update nach 2.60 verwendet. Im Gegensatz dazu erfolgt ein inkrementeller Vergleich zwischen der Belegungswahrscheinlichkeit $p_m(x_m, y_m, z_m)$ einer Zelle m und der Belegungswahrscheinlichkeit $p_m(Z, Z_p)$, die auf Grundlage einer Entfernungsmessung Z_p für m vorgesehen ist. Das Resultat des Vergleiches ergibt sich für die Zelle m wie folgt

$$p_m(x_m, y_m, z_m) = \begin{cases} p_m(Z, Z_p) & \text{falls } p_m = p_{\text{unknown}} \\ p_m(Z, Z_p) & \text{falls } p_m(Z, Z_p) > p_m \\ p_m & \text{sonst.} \end{cases} \quad (3.4)$$

Mit diesem Vorgehen, ist keine Reduktion von $p_m(Z, Z_p)$ für eine Zelle m möglich, was das Freizeichnen verhindert und damit auch die Diskrepanzen unterdrückt, die durch die Diskretisierung entstehen.

Das für ein einzelnes Stereobild generierte M_t wird anschließend mit dem Gitter $M_{1:t}$, welches die Karte aller bis zu diesem Zeitpunkt t aufgezeichneten Daten enthält, fusioniert. Für die Kombination wird in diesem Schritt das Bayes-Update 2.60 verwendet. Bei einer Anzahl von Messungen und deren Positionierung zueinander, wie sie ein Array-Tiefenbild einer Stereokamera bereitstellt, konnte die ungewollte Freizeichnung nicht beobachtet werden.

Für die Sichtbarkeit der ungewollten Freizeichnung ist die Menge der zu einer Punktwolke kombinierten Messung \mathbf{p} und die Überlappung mit vorhergehenden bzw. nachfolgenden Punktwolken ausschlaggebend. Bei der Nutzung eines monogonen LiDAR, wie im Szenario der Abbildung 3.12, ist der Ansatz von Andert [7] auf Basis eines einzelnen Scans keine Lösung. Die ungewollte Freizeichnung findet trotzdem statt. Innerhalb der Punktwolke, die durch die Messungen eines einzelnen Scans S_t aufgebaut wird,

arbeitet der Ansatz von Andert wie vorgesehen, allerdings liegt zwischen den Punktwolken der einzelnen Scans S_t und S_{t+1} nur eine sehr geringe oder keine Überdeckung vor. Darüber hinaus sind die Eintrittswinkel der Laserstrahlen auf die Objektoberfläche recht groß, was zu Freizeichnungen auf Basis der Diskretisierung zwischen den einzelnen Punktwolken führt. Zur Lösung dieses Problems bei monogenen LiDAR-Systemen schlägt Hornung [43] die Pufferung mehrerer Scan S_t in einer Punktwolke $C_{t-n:t} = \bigcup_{t-n}^t S_t$ vor, wobei $n = \text{const}$ und $C_{t-n:t}$ mittels einer FIFO-Umsetzung aktualisiert wird. Der Puffer $C_{t-n:t}$ wird direkt in einen OcTree integriert. Durch die Pufferung wird eine hinreichende Überlappung zwischen den einzelnen Pufferwolken erreicht.

Der Vorteil dieses Vorgehens ist, dass kein zweites Gitter bzw. kein zweiter OcTree benötigt wird. Der Nachteil ist, dass je nach Umfang des Puffers, Messungen mehrfach in den OcTree eingefügt werden, was wiederum einen gesteigerten Prozessierungsaufwand bedeutet.

Darüber hinaus setzt die Methode von Hornung voraus, dass die Sensorposition O_s für eine $C_{t-n:t}$ konstant ist. Bei Verwendung von Punkten \mathbf{p} , welche auf unterschiedlichen O_s basieren, in Kombination mit einer konstanten Sensorposition O_s für $C_{t-n:t}$, ändert sich der Strahlenverlauf zwischen \mathbf{p} und O_s . Das führt zu einer abweichenden Freizeichnung zwischen der gepufferten zur ungepufferten Integration der \mathbf{p} in den OcTree.

3.2.3 Reduktion ungewollter Freizeichnung mittels OcTree

Eine weitere Möglichkeit, die Sichtbarkeit der ungewollten Freizeichnung, die bei der Überführung in eine diskrete Karte entsteht, zu reduzieren, ist die Ausnutzung der Eigenschaften des OcTree. Der OcTree steht hier als Beispiel einer Datenstruktur, die es erlaubt, einen Raum in unterschiedlichen Auflösungen zu diskretisieren.

Das Vorgehen ist dabei wie folgt: die Messungen \mathbf{p} eines Sensors werden einzeln oder als Punktwolke C_i in einen OcTree integriert, was zur Folge hat, dass dessen Blättern Belegtheitswahrscheinlichkeiten $p_m(Z, Z_p)$ zugeordnet werden. Aufbauend auf den Gleichungen 2.62 oder 2.63 werden auch die inneren Knoten mit einer kombinierten $p_m(Z, Z_p)$ der jeweiligen Kind-Knoten ausgestattet.

Die Verwendung von Gleichung 2.63 führt dazu, dass die maximale $p_m(Z, Z_p)$ der Kind-Knoten auf den Elternknoten übertragen wird und damit freie Bereiche mit jeder Reduktion der Baumtiefen d stärker unterdrückt werden. Dieses Verhalten wird ausgenutzt, um die ungewollte Freizeichnung

in ihrer Sichtbarkeit zu reduzieren. Messungen werden auf einer Tiefe d_h in den Baum eingefügt, wobei d_h gleich der Höhe des Baumes ist. Entnommen werden die Messungen allerdings auf einer Tiefe $d_{h-\lambda}$, wobei $\lambda \geq 1$. Das präsentierte Vorgehen ist praktikabel, weist allerdings auch Nachteile auf. Um einerseits eine konstante Auflösung des Baumes für Folgeprozesse bereitzustellen und andererseits eine höhere Auflösung für das Einfügen der Messungen zu ermöglichen, ist eine Steigerung der Baumhöhe um λ notwendig. Pro zusätzlicher Höhenstufe wächst die Anzahl n der Knoten in einem OcTree kubisch. Das bedeutet eine Steigerung der Komplexität θ etwa beim Traversieren $\theta(n)$ durch den Baum. Darüber hinaus sind die Werte für λ Szenarien bedingt und können nicht allgemeingültig festgelegt werden.

3.2.4 Bedingte Freigabe belegter Zelle

Eine weitere Methode, der ungewollten Freizeichnung bei der Diskretisierung der Messdaten entgegenzuwirken, ist, dass die einzelnen Zellen eines Gitters, unabhängig davon, ob es sich um ein Feld oder einen Baum handelt, beim Einfügen einer Messung nicht vollständig unabhängig von ihren Nachbarn betrachtet werden.

Die ungewollte Freizeichnung findet im Allgemeinen zwischen räumlich nahe beieinanderliegenden Messungen statt, siehe Abschnitt 3.2.1. Darüber hinaus tritt die ungewollte Freizeichnung immer zwischen einer als belegt gekennzeichneten Zelle, die geschnitten wird, und einer Zelle auf, die als belegt gekennzeichnet werden soll. Dieser Abschnitt eines Strahls, der zwischen einer Messung \mathbf{p} und der Sensorquelle O_s verläuft, wird genauer betrachtet.

Die ungewollte Freizeichnung tritt in Fällen auf, wenn die Winkel zwischen dem Strahl R_p , dem die Messung \mathbf{p} zugrunde liegt, und der zu vermessenen Oberfläche relativ klein sind. Dieser Zustand kann sich innerhalb eines Gitters in zwei Situationen widerspiegeln [42].

Einerseits ist der Gradient der zu vermessenden Oberfläche lokal etwa orthogonal zu den Achsen des Gitters. In diesem Fall liegen die bereits belegten Zellen, die von der ungewollten Freizeichnung betroffen sind, in direkter Nachbarschaft zu der Zelle in der sich \mathbf{p} befindet. Der Strahl R_p verläuft ab der ersten Berührung mit einer Zelle, für die gilt $p > p_{unknown}$, ausschließlich innerhalb von Zellen mit $p \leq p_{unknown}$. Es werden keine Zellen mit $p \leq p_{unknown}$ geschnitten. Siehe dazu in Abbildung 3.13 die Messung \mathbf{p}_1 .

Verläuft der Oberflächengradient andererseits lokal nicht orthogonal zu den Achsen des Gitters, tritt der beschriebene Alias-Effekt, auch als Trep-

peneffekt bezeichnet, auf. In diesem Fall besteht die Möglichkeit, dass R_p wiederkehrend Zellen mit $p > p_{unknown}$ und $p \leq p_{unknown}$ schneidet. Abbildung 3.13 zeigt diese Situationen mit der Messung p_0 . Um die ungewollte

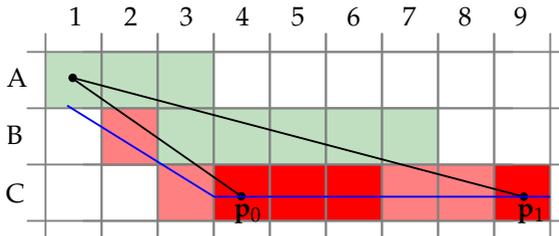


Abbildung 3.13: Strahlenverlauf innerhalb des Belegtheitsgitter, der zur ungewollten Freizeichnung führt.

Freizeichnung in diesen Situationen zu verhindern, wurde eine Methode entwickelt, die es einem Strahl R_p erlaubt, eine gewisse Anzahl th_{occ} von belegten und th_{free} freien Zellen im direkten Umfeld von p zu schneiden, ohne deren Belegtheitswahrscheinlichkeit p um $p(Z, Z_p)$ zu reduzieren. Das setzt voraus, dass die Zellen nicht als unabhängige Messungen betrachtet werden, sondern innerhalb R_p nahe p zueinander in Beziehung gesetzt werden.

Das Referenzieren der Zellen innerhalb R_p wird mittels eines Automaten erreicht, der zusätzlich zu dem verwendeten inversen Sensormodell auf jede Messung angewendet wird.

Der Automat $A = (Q, \Sigma, \delta, \lambda, q_0)$ ist ein endlicher Automat und seine Parameter sind wie folgt definiert; Q Menge von Zuständen, Σ die Zustandsübergangsfunktionen, δ das Eingabealphabet, λ Menge der Endzustände, q_0 der Anfangszustand. Die Eigenschaften und Übergänge von A sind in Gleichung 3.5 und Abbildung 3.14 aufgeführt. Ausgehend von der Zelle m_{O_s} , die O_s enthält, wird im Zustand q_0 gestartet. In diesem Zustand wird mittels σ_0 verblieben, bis entlang R_p die erste Zelle m_k geschnitten wird, welche eine Belegtheitswahrscheinlichkeit $p > p_{unknown}$ aufweist. Darüber hinaus werden in m_{O_s} zwei Laufvariablen $i = 0$ und $j = 1$ initialisiert. Diese beiden Variablen erlauben eine flexible Steuerung der Toleranzen th_{occ} und th_{free} bei der bedingten Freizeichnung. Die Toleranz th_{occ} gibt einerseits an, wie viele als belegt gekennzeichnete Zellen sich direkt hintereinander befinden dürfen, bevor die Zelle m_p folgt, welche p enthält. Das ermöglicht ein bedingtes Freizeichnen, bei einem lokal zu den Gitterachsen orthogonal verlaufenden Oberflächengradienten. Andererseits legt th_{free} die maximale Anzahl der als frei gekennzeichneten Zellen fest, die zwischen zwei be-

$$\begin{aligned}
 Q &= \{q_0; q_1; q_2; q_3; q_4\} \\
 \Sigma &= \left. \begin{cases} \sigma_0 = \{p_m \leq p_{unknown} \ \& \ i \leq th_{free}\}; \\ \sigma_1 = \{p_m \leq p_{unknown} \ \& \ i > th_{free}\}; \\ \sigma_2 = \{p_m > p_{unknown} \ \& \ j \leq th_{occ}\}; \\ \sigma_3 = \{p_m > p_{unknown} \ \& \ j > th_{occ}\}; \\ \sigma_4 = \{p_m > p_{unknown} \ \& \ m = \lfloor \frac{p_i - t_{mi}}{d_m} + 0.5 \rfloor\} \end{cases} \right\} \quad (3.5) \\
 \lambda &= \{q_3; q_4\}
 \end{aligned}$$

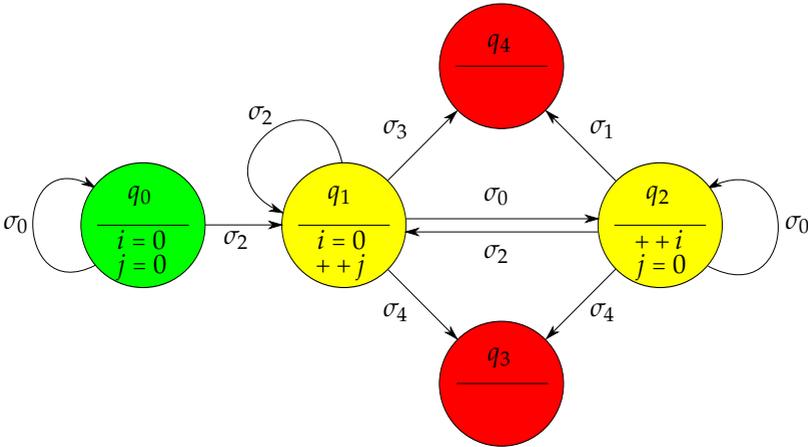


Abbildung 3.14: Endlicher Automat zur Umsetzung der bedingten Freizeichnung innerhalb eines Strahls R_p . Nach [42].

legten Zellen geschnitten werden dürfen. Dies ermöglicht eine Beeinflussung der ungewünschten Freizeichnung beim Alias-Effekt. Wachsen i und j über die jeweilige Toleranz hinaus, wird mit dem klassischen, unbedingten Verfahren bei der Integration einer Messung in das Gitter fortgefahren. Das Erreichen einer belegten Zelle $m_{occ,k}$, $k = 1$ ermöglicht einen Übergang σ_2 nach q_1 . Die Belegtheitswahrscheinlichkeit von $m_{occ,1}$ wird in σ_1 nicht angepasst. Sie wird aber inklusive des Wertes $p(Z, Z_p)$, der sie verändern sollte, gepuffert, um später darauf zurückzugreifen. Basierend auf q_1 sind vier Übergänge möglich.

Folgt auf m_1 die Zelle m_p ist das Ende des Strahls erreicht und es wird mittels σ_4 der Endzustand q_3 erreicht. Das Erreichen von q_3 führt dazu, dass die bedingte Freizeichnung umgesetzt wird. Das bedeutet, dass der Puffer mit den zurückgehaltenen Aktualisierungen geleert wird, ohne dass diese angewendet werden. Es wird ausschließlich die Zelle m_p mittels Gleichung 2.60 aktualisiert.

Folgt auf $m_{occ,1}$ wiederum eine belegte Zelle $m_{occ,2}$ wird unter der Bedingung $j \leq th$ in q_1 verbleiben. In diesem Fall wird $++j$ angewendet. Wenn die Bedingung $j \leq th$ verletzt wird, wird der Übergang nach q_4 gewählt. Der Zustand q_4 ist ebenfalls ein Endzustand und zeigt an, dass die bedingte Freizeichnung abgebrochen wurde. In diesem Fall werden alle im Puffer befindlichen Aktualisierungen umgesetzt und ab der aktuellen Zelle m_k wird mit dem bekannten Bayes-Update 2.60 fortgefahren.

Das mehrmalige Erreichen von q_1 mittels σ_2 zeigt an, dass es sich um die j -belegte Zelle handelt, die ohne Unterbrechung erreicht wurde. Diese Schleife des Automaten wird zur Auflösung des Szenarios eingesetzt, welches zum Beispiel bei p_1 in Abbildung 3.13 entsteht. Um nicht unbeabsichtigt dynamische Hindernisse in der Karte zu belassen, die sich direkt vor statischen Objekten befanden, wird die Bedingung $j \leq th_{occ}$ verwendet.

Folgt auf $m_{occ,1}$ eine frei Zelle $m_{free,1}$ wird mittels σ_0 nach q_2 gewechselt. Dieser Zweig des Automaten dient zur Behandlung der ungewollten Freizeichnung bei Alias-Situationen, vergleichbar zu der Situation von Messung p_0 in Abbildung 3.13. Folgt auf q_2 eine weitere $m_{free,1}$, wird unter der Bedingung $i \leq th$, in q_2 verblieben. Wird $i \leq th_{free}$ in q_2 verletzt, wird nach q_4 gewechselt und alle gepufferten und zukünftigen Aktualisierungen mittels normalem Bayes-Update umgesetzt.

Folgt auf $m_{free,1}$ oder eine seiner Folgezellen, welche die Bedingung $i \leq th_{free}$ nicht verletzen, eine belegte Zelle $m_{occ,2}$, wird zurück zum Zustand q_1 gewechselt und die bedingte Freizeichnung zur Behandlung des Alias-Effekts neu initialisiert. Das ermöglicht auch die Behandlung längerer Treppenbereiche.

Folgt auf $m_{free,1}$ oder eine seiner Folgezellen, welche die Bedingung $i \leq$

th_{free} nicht verletzen, die Zelle m_p wird der Endzustand q_3 erreicht und die bedingte Freizeichnung umgesetzt.

3.2.5 Validierung

Für das vorgestellte Problem der Abbildung eines Umweltszenarios durch eine diskrete Karte unter Verwendung eines monogonen LiDAR (Abschnitt 3.2) werden die in den letzten drei Abschnitten vorgestellten Lösungsansätze gegenübergestellt.

Der erste Ansatz 3.2.3 ist das Integrieren von Messungen bei einer höheren Auflösung in das kartierende Medium und die Entnahme bei einer geringeren Auflösung. Für eine vereinfachte Bezeichnung wird dieser Ansatz im Verlauf des aktuellen Abschnittes mit A bezeichnet.

Der zweite Ansatz, der auch von Hroning [43] vorgestellt wird und für eine Gruppe von Ansätzen steht, kombiniert mehrere Scans S_i im Vorfeld bevor die Vereinigungsmenge anschließend in das kartierende Medium, hier einen OcTree, integriert werden. Dieser Ansatz wird mit B bezeichnet. Die dritte Umsetzung, die bedingte Betrachtung von geschnitten Gitterzellen entlang eines Messstrahls (Abschnitt: 3.2.4) wird mit C betitelt.

Die Bewertung erfolgt in zwei Schritten. Zuerst wird auf die ausgewählten Parametersets pro Ansatz eingegangen. Die Validierung wird dabei visuell vorgenommen. Ziel ist es, eine geschlossene Detektion des Bodens zu erreichen.

Aufbauend auf den dargelegten Sets wird anschließend eine Performanzanalyse hinsichtlich Laufzeit, Komplexität und Speicherbedarf durchgeführt.

Als Beispielszenario für die Bewertung wird ein Überflug (Abbildung 3.12) eines unbemannten Hubschraubers über ein Feld mit zwei künstlichen Hindernissen, zwei Torstangen, verwendet, wobei das UA ein LiDAR zur Datengenerierung transportiert. Zur Aufzeichnung der Daten wurde der „Hokuyo UTM-30LX“ mit einer Anbauposition vergleichbar zu Abbildung 4.25 verwendet.

Parametersets für geschlossene Bodenrepräsentation

Die Ausgangssituation, das Einfügen der aufgezeichneten Messungen in einen OcTree mit Zellabmessung $d_m = 0,5 m$ und eine Baumhöhe $h = 10$, ohne jegliche Anwendung einer Erweiterung zeigt Abbildung 3.15. Die Färbungen der dargestellten Voxel spiegeln die Belegungswahrscheinlichkeit $p(Z)$ wieder. Die Messungen des Lasers werden grundsätzlich mit $p(Z) = 0.8$ in den OcTree integriert. Abweichende Belegungswahrscheinlichkeiten sind in der Abbildung durch andere Farben dargestellt.

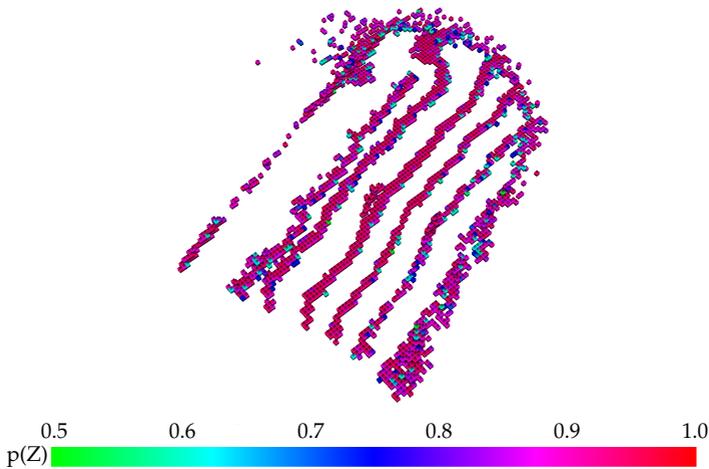


Abbildung 3.15: Ausgangssituation, ohne Berichtigung der ungewollten Freizeichnung.

lichkeiten weisen auf eine Anwendung eines Bayes-Update hin. Die Darstellungen zeigen nur die belegten Voxel, was Belegungswahrscheinlichkeiten $0.5 \leq p(Z) \leq 1.0$ entspricht.

Die Messungen des monogonen LiDAR wurden in einen OcTree mit $d_m = 0,5 m$ und $h = 10$ eingefügt und auf der gleichen Höhe wieder entnommen. Es ist zu erkennen, dass große Teile der Bodenplatte nicht durch Voxel repräsentiert werden. Darüber hinaus enthält die Abbildung eine große Anzahl von Voxeln, die eine Belegtheitswahrscheinlichkeit $p(Z) \leq 0.8$ aufweisen, was auf eine gegensätzliche Interpretation der Belegung zurückzuführen ist. Eine Ursache neben Rauschen ist etwa die Diskretisierung des Raumes, wie es in Abschnitt 3.2 bzw. den Abbildungen 3.11 verdeutlicht wurde.

Ziel der drei Ansätze ist es, diese Lücken zu schließen bzw. die Belegungswahrscheinlichkeiten der Voxel so zu manipulieren, dass die ungewollte Freizeichnung, welche durch die Diskretisierung entsteht, nicht sichtbar wird. Um dies zu erreichen, müssen die Ansätze jeweils unterschiedlich parametrisiert werden. Eine für das gewählte Szenario bestimmte Parameterauswahl zeigt Tabelle 3.1.

Ansatz A: Die Anwendung des Ansatzes A beim Einfügen der Messungen in den OcTree auf einer größeren Tiefe $d = 11$ bei A(2) und $d = 12$ bei A(3) zeigen die Abbildungen 3.16a bzw. 3.16b. Bereits beim Einfügen der

Ansatz		Parametersets			
Nr.	Beschreibung	Set 1	Set 2	Set 3	Set 4
A	Einfügen bei Tiefe d_{in} , Entnahme bei Tiefe d_{out} , wobei $d_{out} < d_{in}$	$d_{in} = 12$ $d_{out} = 11$	$d_{in} = 12$ $d_{out} = 10$		
B	Kombination mehrerer Scans S_i zu einer Punktwolke $C = \bigcup_{i=0}^j S_i$	$j = 20$	$j = 40$	$j = 60$	$j = 80$
C	Einfügen mittels bedingter Freizeichnung	$th_{free} = 1$ $th_{occ} = 1$	$th_{free} = 4$ $th_{occ} = 1$	$th_{free} = 1$ $th_{occ} = 4$	$th_{free} = 2$ $th_{occ} = 8$

Tabelle 3.1: Parameter zur Unterdrückung der ungewollten Freizeichnung.

Messungen auf einer Tiefe d , die um eins höher ist als die Tiefe auf der der Baum verwendet werden soll, zeigt eine Reduzierung der ungewollten Freizeichnung. Das Einfügen auf der Tiefe, die zwei Ebenen über der zur Verwendung beabsichtigten Tiefe liegt, zeigt augenscheinlich eine nahezu vollständige Unterdrückung der durch die Diskretisierung auftretenden Problematik. Allerdings enthält auch diese Umsetzung kleine Löcher in der Größe von einzelnen Voxeln.

Ansatz B: Die Abbildung 3.17 präsentiert das Resultat des Ansatzes B bei dem jeweils j -Scans des monogonen Lasers zu einer Punktwolke kombiniert werden und anschließend in den OcTree integriert werden. Bei der Steigerung von B(1) mit $j = 20$ hinzu B(4) mit $j = 80$ ist eine Reduktion der Freiräume zu erkennen. Darüber hinaus ist ein Rückgang der Voxel zu erkennen, die ein $p(Z) \leq 0.8$ aufweisen, was ebenfalls den Rückgang von widersprüchlichen Belegungen verdeutlicht. Simultan steigt augenscheinlich die Anzahl der Voxel mit $p(Z) \geq 0.8$. Was für eine gegenseitige Stützung der Messungen untereinander spricht. Das Kombinieren mehrerer Scans führt eindeutig zu einer Reduktion der Sichtbarkeit der Effekte, die durch

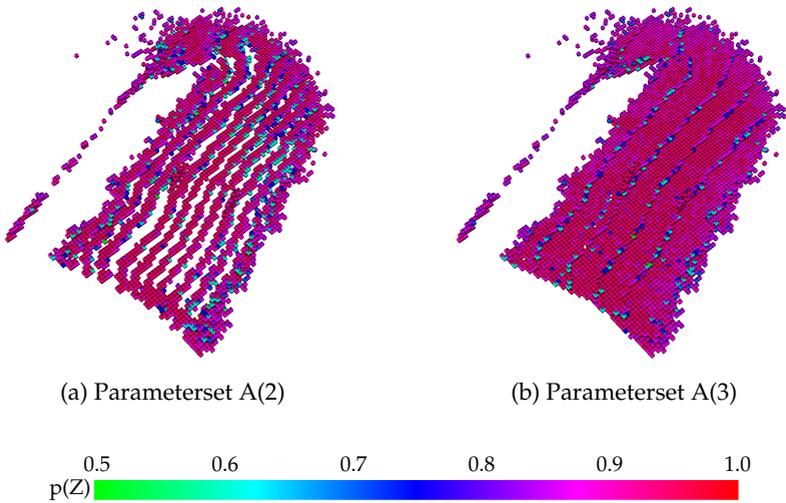


Abbildung 3.16: Ansatz A für das Beispiel einer planaren Bodenplatte

die Diskretisierung auftreten. Allerdings ist selbst bei B(4), wie Abbildung 3.17d zeigt, keine vollständige Lösung der Problematik möglich.

Ansatz C: Die Abbildung 3.18 präsentiert das Resultat des Ansatzes C bei dem jeweils th_{occ} belegte und th_{free} freie Voxel in ihrer Belegungswahrscheinlichkeit nicht beeinflusst werden, wenn sie von Strahlen geschnitten werden. Die Abbildungen 3.18a und 3.18b zeigen die Resultate der Parametersets C(1) und C(2) und spiegeln bis auf marginale Änderungen die Ausgangssituation in Abbildung 3.15 wider. Eine Erhöhung von th_{free} in C(3) führt allerdings zu einer Reduktion der ungewollten Freizeichnung. Die Anwendung von C(4) in Abbildung 3.18d zeigt schlussendlich eine wiederum nahezu vollständige Reduktion der ungewollten Freizeichnung, auch wenn noch Voxel mit einer $p(z) \leq 0.8$ zu erkennen sind.

Diskussion Die Ergebnisse der jeweiligen Parametersets zeigen, dass mit allen drei Ansätzen eine geschlossen Bodenrepräsentation erzeugt werden kann. Bei der visuellen Gegenüberstellung aller Parametersets zeigen A(3), B(4) und C(4) für jeden Ansatz die vermeintlich besten Ergebnisse mit einer nahezu vollständig geschlossen Bodenplatte mit den beiden Hindernissen. Diese drei Parametersets werden im Folgenden einer weiteren Betrachtung

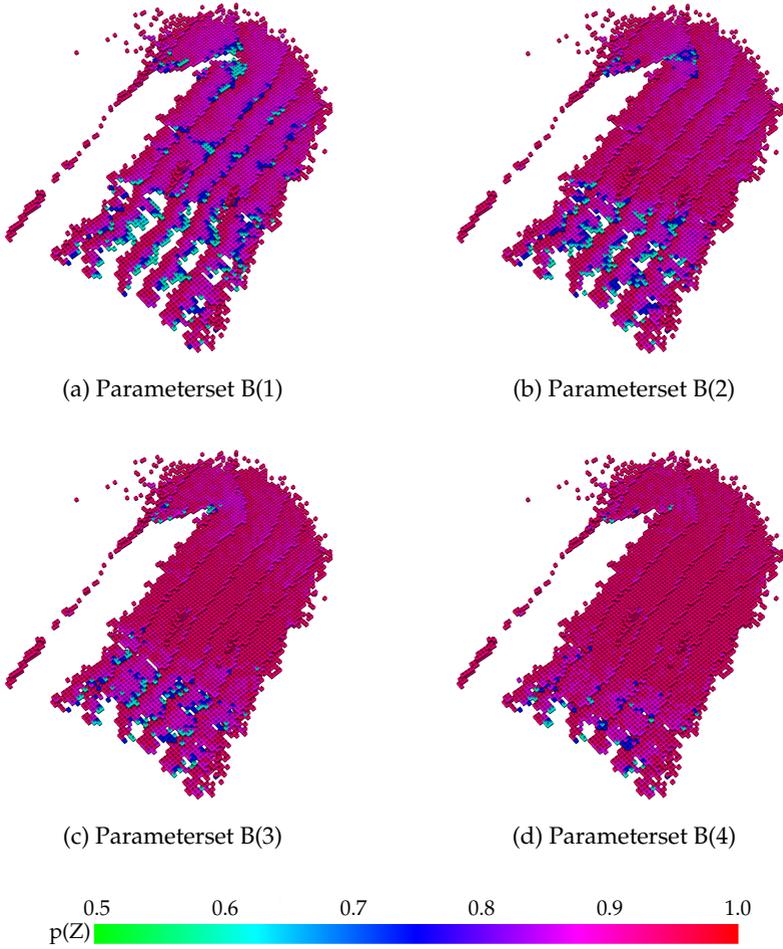


Abbildung 3.17: Ansatz B für das Beispiel einer planaren Bodenplatte

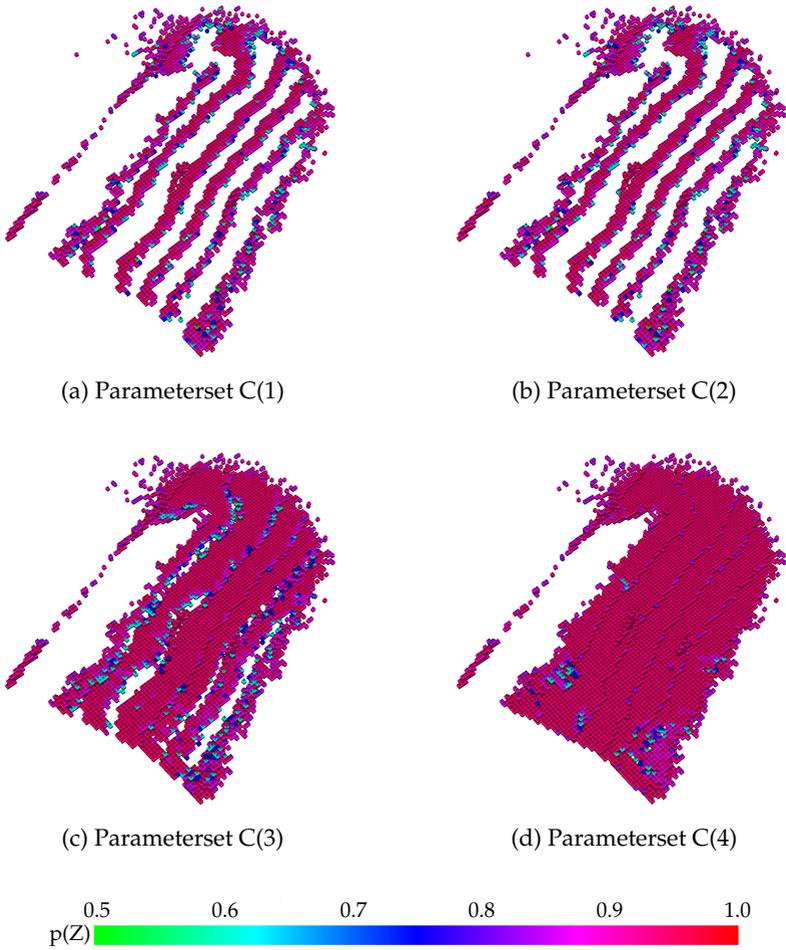


Abbildung 3.18: Ansatz C für das Beispiel einer planaren Bodenplatte

unterzogen, um eine vergleichende Performanzanalyse zwischen den drei Ansätze zu ermöglichen.

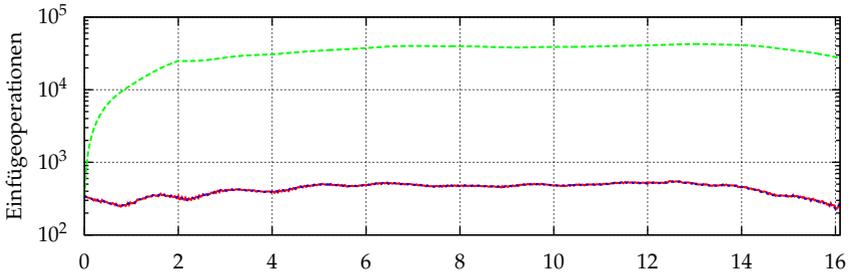
Performanzanalyse

Die dargestellten Ansätze werden in Online-Anwendungen eingesetzt. Neben einer möglichst korrekten Abbildung der vermessenen Umwelt müssen sie daher auch auf der zur Verfügung stehenden Hardware¹ an Bord des UA berechenbar sein. Um benötigte Rechnerperformanz der jeweiligen Ansätze bewertbar zu machen, wurden Größen wie Laufzeit pro einzufügender Punktwolke, Integrationsschritte in den OcTree pro einzufügender Punktwolke und Speicherbedarf des OcTree für die ausgewählten Parametersets A(3), B(4) und C(4) bestimmt. Die Resultate zeigt Abbildung 3.19. Bei der Gegenüberstellung der Anzahl der Einfügeoperationen pro Punktwolke zwischen A(3), B(4) und C(4) ist zu erkennen, dass B(4) von A(3) und C(4) abweicht. Diese Abweichung ist auf die Kombination von bis zu 80 Scans zu einer Punktwolke zurückzuführen. Nach einer Aufbauphase über die ersten 80 Scans entspricht die Anzahl der Einfügeoperationen pro Punktwolke von B(4) etwa dem achtzig-fachen von A(3) und C(4).

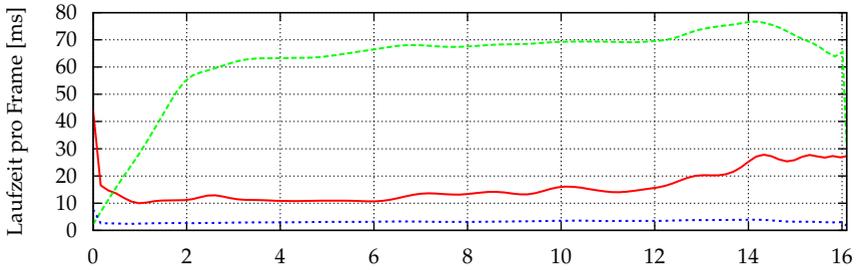
Die große Anzahl von Einfügeoperationen bei B(4) wirkt sich ebenfalls auf die benötigte Laufzeit pro Punktwolke aus. Allerdings ist die Differenz zu den Resultaten von A(3) und C(4) in diesem Fall nicht so groß wie bei der Anzahl der Einfügeoperationen. Dieses Resultat ist auf das Zusammenfassen von Messungen zu einer Punktwolke zurückzuführen, was wiederum dazu führt, dass nicht alle Messungen direkt in den OcTree integriert werden. Liegen etwa mehrere Messungen räumlich in einem Voxel, ist es ausschließlich notwendig, die Aktualisierung des Voxel einmal durchzuführen. Folgeoperationen wie etwa das Traversieren zur Wurzel, um den gesetzten Status an die Eltern zu übertragen, ist in diesem Fall im Gegensatz zur Integration jeder Einzelmessung ebenfalls nur einmal notwendig, was zu einer Reduzierung der benötigten Zeit führt. Im Großen und Ganzen führt die große Anzahl der zu integrierenden Messungen in den OcTree allerdings zu einem starken Anwachsen der benötigten Laufzeit für den Ansatz B(4) gegenüber A(3) und C(4).

Bei der Anzahl der Einfügeoperationen pro Punktwolke sind die Ansätze A(3) und C(4) etwa identisch, was darauf zurückzuführen ist, dass jede Messung direkt vom Laser kommt und in den OcTree übertragen wird, ohne dass sie mit weiteren Messungen kombiniert oder verworfen wird. Bei der Betrachtung der Laufzeit pro Punktwolke ist allerdings ein Unterschied zu erkennen, der auf die Größe des OcTrees und schlussendlich auf

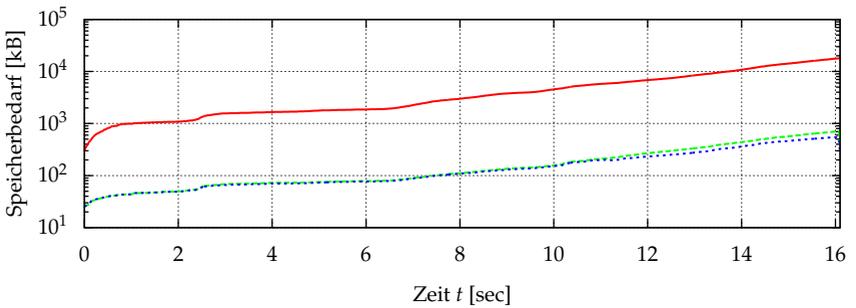
¹Verwendeter Rechner: i5-4210(1.7 bis 2.7GHz), 12GB RAM



(a) Anzahl der Einfügeoperationen in den OcTree pro Punktwolke.



(b) Laufzeit pro Punktwolke.



(c) Speicherbedarf des OcTree.

Abbildung 3.19: Gegenüberstellung der Parametersets A(3), B(4) und C(4).

die Anzahl der Folgeoperationen bei der Messwertintegration in den Oc-Tree zurückzuführen ist.

Ein OcTree $N = 8$ mit der Höhe $h = 10$ weist nach

$$n = \frac{N^{h+1} - 1}{N - 1} \quad (3.6)$$

eine Gesamtzellenanzahl von $n = 1.227.133.513$ Zellen auf. Bei $h = 12$ entspricht das $n = 78.536.544.841$ Zellen. Das asymptotische Verhalten von Operationen auf dem OcTree, wie etwa der Zugriff auf Zellen oder die Suche ($\mathcal{O}(\log n)$) und das Traversieren ($\Theta(n)$), sind durch n bedingt. Folglich steigt mit zunehmendem n auch die benötigte Anzahl an Iterationen, die etwa zur Aktualisierung der Elternknoten benötigt wird und damit die dafür benötigte Zeit pro Messung.

Aufbauend auf der höheren Auflösung, mit der die Messungen in den Oc-Tree bei A(3) integriert werden, folgt, dass weniger Messungen in Voxel kombiniert werden und mehrere Voxel auf einzelnen Messungen beruhen. Der Speicherbedarf eines Voxels ist konstant und unabhängig davon, auf welcher Tiefe sich der Voxel befindet. Mit zunehmender Tiefe steigt allerdings die Anzahl der inneren Knoten, welche die Blätter mit der Wurzel verbinden. Die Kombination aus einer größeren Anzahl von Blättern und inneren Knoten führt daher zu einem erhöhten Speicherbedarf. Dies spiegelt auch die Gegenüberstellung des für den OcTree benötigten Speichers der jeweiligen Ansätze wider. Zu erkennen ist der erhöhte Speicherbedarf von A(3).

Der Ansatz B(3) weist gegenüber C(4) ebenfalls einen leicht erhöhten Speicherbedarf auf. Beide Ansätze integrieren die Messungen auf $h = 10$ in den OcTree und enthalten folglich etwa die gleiche Anzahl von belegten Blättern (4880 Voxel bei B(3) und 4895 Voxel bei C(4)) und belegten inneren Knoten am Ende des Szenarios (Scan 645). Diese Differenz lässt sich mit den als belegt deklarierten Knoten nicht deuten. Der Unterschied erklärt sich aus der Anzahl der als frei eingeordneten Knoten. Bei B(3) handelt es sich am Ende des Überfluges um 20957 Voxel und bei C(4) lediglich um 14590 Voxel.

Dieser Unterschied ist auf das Einfügen mittels einer Punktwolke zurückzuführen, für welche ein einziger Sensorursprung angenommen wird. Für die Berechnung der Hits einer Punktwolke wird die zum Zeitpunkt des Hits vorliegende Sensorposition verwendet. Zur Bestimmung der Freizeichnung wird allerdings nur die Sensorposition des letzten Hits verwendet und für alle Hits der Punktwolke übernommen. Dies führt zur Freizeichnung von Voxeln, die bei einer korrekten Zuordnung von Sensorposition und Hit nicht aktualisiert wurden. Eine Verdeutlichung dieses Sachverhal-

tes zeigt Abbildung 3.20. Die schwarzen durchgehenden Strahlen verlau-

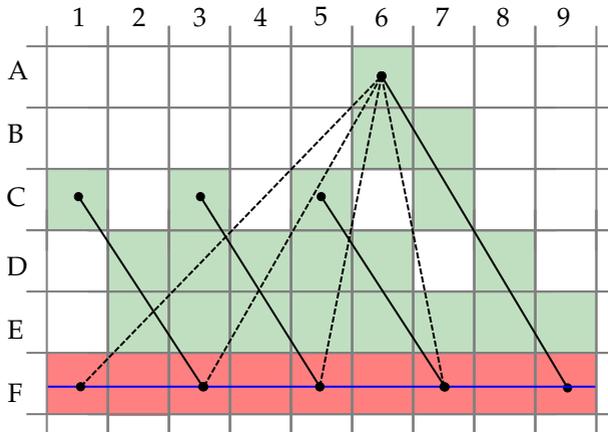


Abbildung 3.20: Direktes Einfügen einzelner Scans gegenüber einem kombinierten Einfügen mehrerer Scans.

fen von der Sensorposition zum Hit auf der blauen Hindernisoberfläche. Die Voxel, welche von diesen Strahlen geschnitten werden, werden frei gezeichnet, hier durch die Grünfärbung dargestellt. Der Voxel, welcher den Hit trifft, wird als belegt, hier rot, markiert. Dieses grundlegende Prinzip trifft auch auf die Punktwolke zu, für die allerdings nur die letzte Sensorposition (A6) zur Freizeichnung herangezogen wird. Durch dieses Vorgehen werden auch Voxel (B5, C4, C6 und D7) von Strahlen (gestrichelt) geschnitten, die zuvor keine Aktualisierung erfuhren, was zu dem Anstieg der als frei deklarierten Voxel führt. Als weitere Folge dieses Vorgangs wird Raum, für dessen Querung eigentlich keine Messungen vorliegen, als frei deklariert, was beim tatsächlichen Vorhandensein eines Objektes zu einer Kollision führen kann.

Fazit

Resultierend auf der gezeigten visuellen und numerischen Gegenüberstellung der drei Ansätze A, B und C sowie im Speziellen der Parametersets A(3), B(4) und C(4) wurde das Parameterset C(4) für das hier präsentierte Szenario, einer überflogenen Bodenplatte, als der am geeignetste Ansatz zur Integration von monogonen LiDAR-Daten bestimmt. Für dieses Urteil waren die geringe Laufzeit pro Punktwolke, der geringe Speicherbedarf und die respektable Darstellung der vermessenen Umwelt ausschlagge-

bend. Aufbauend auf diesem Ergebnis werden die weiteren Auswertungen unter Berücksichtigung des Parametersets C(4) durchgeführt. Bei Abweichenden Parametereinstellungen wird auf diese separat hingewiesen.

Kapitel 4

Erweiterung einer monogonen Lasersicht mittels Kamera

Für den automatisierten Betrieb von kleinen unbemannten Hubschraubern in einer unbekanntem Hinderniskulisse, etwa unter Randbedingungen wie dem Fliegen außerhalb der Sicht des Operators oder bei Kommunikationsverlust, ist an Bord des unbemannten Systems eine dreidimensionale Detektion der Hindernisse erforderlich, um aufbauend auf den ermittelten Hindernissen ein Ausweichen zu ermöglichen. Ein unbemannter Hubschrauber kann sich nahezu unabhängig dreidimensional bewegen. Um das Bewegungspotenzial des Hubschraubers nicht zu limitieren, wird ebenfalls eine dreidimensionale Wahrnehmung der Umwelt angestrebt, um auf Basis der gewonnenen Daten auch ein sicheres dreidimensionales Ausweichen von Hindernissen zu ermöglichen. Im Rahmen dieser Arbeit wird unter der dreidimensionalen Detektion eines Hindernisses die Detektion des Raumes verstanden, der von dem Hindernis mindestens teilweise belegt wird. Um welche Art von Hindernis (Baum, Haus, Mensch, ...) es sich dabei handelt, ist nicht von Interesse.

Einführend wird ein Überblick über bestehende Ansätze gegeben, welche verschiedene LiDAR und Kamera-Konfigurationen einsetzen, um auf Basis beschränkter Sensordaten eine umfassende Sicht auf die Umwelt zu ermöglichen. Darauf aufbauend werden das Konzept und der Ablauf des hier vorgestellten Ansatzes dargelegt.

4.1 Stand der Technik

Auf Basis der gegebenen Randbedingungen des UA (Abschnitt 4.6.1) und zur Verfügung stehenden Sensoren (Abschnitt 2.3) werden Kameras und LiDAR als Ausgangspunkt für die Hindernisdetektion betrachtet. Sensoren wie Sonar oder Radar wurden aufgrund ihrer Leistungsspezifika nicht weiter berücksichtigt. Ein Überblick über bestehende Systeme oder Verfah-

ren, der keinen Anspruch auf Vollständigkeit hat, wird im Folgenden präsentiert und im Hinblick auf die angestrebte Anwendung in einem kleinen UA bewertet. Es wird in homogene und heterogene Systeme unterschieden, wobei sich die Unterteilung auf die Art der Sensoren bezieht.

4.1.1 Homogene Konfigurationen

LiDAR kombinieren mehrere seriell erzeugte Messungen zu einem Abbild der Umwelt. Das serielle Abtasten der Umwelt wird mit einer Mechanik erreicht, die je nach Umfang des angestrebten Sichtfeldes in Größe und Umfang variiert. Scanner, die ein zweidimensionales Sichtfeld vergleichbar dem einer Kamera¹ aufspannen, können oft aufgrund ihres Gewichts und ihrer Größe [80, 96] von kleinen UA nicht transportiert werden.

Scanner, die Messungen ausschließlich für eine Schnittebene generieren, werden auch als monogone LiDAR bezeichnet. Sie können von kleinen UA transportiert werden, geben aber keine Information über den Zustand der Umwelt über- und unterhalb der Ebene ab [86].

Ansätze [96, 97], monogone LiDAR mittels zusätzlicher externer Antriebe und Spiegel zu erweitern und darauf aufbauend eine dreidimensionale Abtastung der Umwelt zu erreichen, haben zwei Schwachpunkte. Einerseits bringen die Erweiterungen zusätzliches Gewicht mit in die Sensorconfiguration, was oft zu der gleichen Problematik wie bei einem dreidimensional scannenden System führt. Sie sind zu schwer für kleine UA. Andererseits ist die Aktualisierungsrate, mit der das erweiterte Sichtfeld abgetastet werden kann, durch die Aktualisierungsrate des Ausgangssystems bedingt, was zu sehr geringen Aktualisierungsraten des Gesamtsystems führen kann. Geringe Aktualisierungsraten können in einer dynamischen Umwelt oder einem hochdynamischen System, wie einem UA, zu Problemen führen.

Um dennoch einen Betrieb von LiDAR an kleinen UA sowie eine dreidimensionale Detektion der Umwelt zu ermöglichen, nutzen [36, 65] die Eigenbewegung des UA. Merz [65] prägt dem UA eine zusätzliche zyklische Lageänderung auf, wodurch eine zusätzliche Dimension für Messungen aufgespannt wird. Die aufgeprägte Bewegung führt allerdings zu einer Einschränkung des Flugverhaltens, was den Einsatz für höhere Aufgaben² beschränkt. Grzonka [36] nutzt ausschließlich die Eigenbewegung des UA

¹Kamera mit einem zweidimensionalen CCD und einer entsprechenden Optik (z.B. $f=5\text{mm}$), die es erlaubt, zweidimensionale Aufnahmen der Umwelt zu erzeugen.

²Das Fliegen an sich, Pfade abfliegen oder Hindernisse detektieren und vermeiden ist nicht die eigentliche Zielstellung eines UA. Diese Aktivitäten sind allein Voraussetzungen, um höhere Aufgaben wie die Inspektion, Suchen oder den Transport von Gütern zu ermöglichen.

bei Richtungs- sowie Positionsänderungen aus. Dieses Vorgehen setzt eine kontinuierliche Ausgleichsbewegung um mindestens eine Achse des UA voraus. Ist diese Ausgleichsbewegung, etwa bei einem kontinuierlichen, unbeschleunigten Flug innerhalb der Laserebene, nicht gegeben, ist es ebenfalls notwendig eine künstliche Bewegung anzulegen.

Kameras ob nun monokular oder auch multiokular werden ebenfalls zur dreidimensionalen Hindernisdetektion eingesetzt. Eine weit verbreitete Variation eines multiokularen Ansatzes ist der Stereokameraaufbau [7, 45]. Bei der Stereoskopie wird mittels korrespondierende Bildpunkte der beiden Kamerabilder eine Triangulation durchgeführt, deren Ergebnis eine indirekte Entfernungsmessung ist.

Eine indirekte Entfernungsmessung ist ebenfalls mit einer monokularen Kamera möglich, wenn die extrinsische Transformation zwischen zwei Aufnahmen bekannt ist und sich die Sicht der Aufnahmen ebenso wie beim multiokularen Aufbau hinreichend überlappt.

Die indirekte Entfernungsmessung enthält systematische Unzulänglichkeiten, die zu einer quadratisch abnehmenden Entfernungsauflösung mit zunehmender Entfernung führen. Weiteres Fehlerpotenzial ist durch inkorrekte Korrespondenzzuweisungen, die bspw. durch eine dynamische Umwelt, homogene Oberflächen oder Verformungen des Stereokameraaufbaus gegeben und kann zu falschen Entfernungsmessungen führen. Basierend auf den Mängeln wird den Entfernungsmessungen, welche mittels Stereoskopie entstanden sind, ein geringeres Vertrauen als bspw. Lasermessungen zugesprochen [46].

4.1.2 Heterogene Konfigurationen

Stereokamera und monogoner LiDAR: Eine Gegenüberstellung der Vor- und Nachteile eines monogonen LiDAR und einer Stereokamera bei der Hindernisdetektion für kleine UA präsentiert u.a. Hrabar [46]. Die Gegenüberstellung betrachtet neben einer jeweiligen Umsetzung mit den einzelnen Sensoren auch eine Fusion beider Sensoren. Die Fusion der unabhängigen Sensoren erlaubt auf der einen Seite eine Kombination der Sichtfelder und auf der anderen Seite eine Steigerung des Vertrauens bzw. der Messgenauigkeit. Eine derartige Fusion präsentiert neben Hrabar auch Moghadam [66] für Bodenfahrzeuge. Beide Ansätze fusionieren 2D-LiDAR- mit den Stereokameradaten in einem dreidimensionalen probabilistischen Belegtheitsgitter.

Bei seiner Gegenüberstellung kommt Hrabar zu dem Urteil, dass die Stereokamera sowohl als homogene Konfiguration als auch als Komponente

in einer Fusion für die Hindernisdetektion außerhalb von Gebäuden ungeeignet ist und eine homogene Laserkonfiguration vorzuziehen ist.

Hrabar verwendet eine Central Level Fusion mittels Belegtheitsgitter, in der die Sensordaten ohne eine Vorverarbeitung integriert werden. Fehlmessungen der Stereokamera können durch den LiDAR im Gitter nur dann korrigiert werden, wenn sich zu korrigierende Messungen mindestens einmal innerhalb des Sichtfeldes des LiDAR befinden. Das Sichtfeld des Lasers ist allerdings kleiner als das der Stereokamera, so dass nicht davon ausgegangen werden kann, dass jede Stereokameramessung durch eine Lasermessung beeinflusst wird. Ein Teil der Stereodaten unterliegt keiner Fusion mit Laserdaten und geht ungefiltert in die Hindernisrepräsentation ein. Das Resultat ist, dass die Nachteile der Stereokamera wie Phantomhindernisse [7] ebenfalls im Output der Fusion enthalten sind.

Moghadam [66] generiert seine Resultate vor allem innerhalb von Gebäuden bei einer statischen Umwelt und Beleuchtung und erreicht erwartungsgemäß bessere Resultate mit der Stereokamera als Hrabar. Aber auch in dieser Umsetzung weisen die Stereokameradaten ein hohes Rauschen und falsche Entfernungsmessungen auf. Der Aussage von Hrabar, allein auf den LiDAR zu vertrauen, widerspricht Moghadam, da Hindernisse die über- bzw. unterhalb der Scanebene liegen nicht detektiert werden [86]. Moghadam kommt zu dem Ergebnis, dass der alleinige Einsatz von einem 2D-LiDAR oder einer Stereokamera nicht vertretbar ist und nur eine Fusion hinreichend genau ist, um verlässliche Daten für eine Hindernisdetektion bereitzustellen.

Monokulare Kamera und monogoner LiDAR: Die Stereoskopie weist Nachteile auf, die einen Einsatz für die Hindernisdetektion bzw. die Generierung von Tiefendaten außerhalb von Gebäuden erschweren. Es gibt Ansätze, welche den direkten Einsatz der Stereoskopie bei der Hindernisdetektion umgehen. Ein derartiger Ansatz wurde von Baltzakis [11] vorgestellt, wobei die Stereoskopie nicht direkt zur Hindernisdetektion verwendet wird, sondern um die Hindernisdaten, welche mittels eines monogonen LiDAR erstellt wurden, aber außerhalb dessen Sichtfeldes liegen, zu validieren.

Die alleinige Verwendung eines monogonen LiDAR vernachlässigt Überhänge, die sich über- bzw. unterhalb der Scanebene befinden. Zur Identifizierung dieser Objekte nutzt Baltzakis eine monokulare Kamera und einen stereoskopischen Ansatz. Das Einbeziehen des stereoskopischen Ansatzes in eine Fusion führt ebenso zur Inklusion von dessen Nachteilen. Hinzu kommt, dass kein statischer Stereoaufbau verwendet wird.

Die Herleitung der Stereoinformation aus den zeitlich versetzt aufgezeichneten Kamerabildern setzt voraus, dass die Transformation zwischen den Aufnahmezeitpunkten kontinuierlich bestimmt wird. Dieses kann bspw. mit Odometrie oder optischer Navigation erfolgen. Unzulänglichkeiten des jeweils eingesetzten Verfahrens sind weitere Fehlerquellen, welche die Aussagekraft dieses stereoskopischen Ansatzes weiter reduzieren. Ein weiteres Problem bei einem monokularen Ansatz ist die Bereitstellung stereoskopischer Information entlang der optischen Achse bei geringer Änderung des Blickwinkels, da die Verschiebung der Bildfeatures zu gering ist, um diese hinreichend genau abzuleiten [45, 47]

Ansätze zur Hindernisdetektion, die eine Fusion aus monogonen LiDAR- und monokularen Kameradaten ohne die Ausnutzung von stereoskopischen Verfahren verwenden, werden von Peynot [73] und Garcia-Alegre et al. [31] vorgestellt. Peynot verwendet die Intensitätsbilder der Kamera, um die Entfernungsdaten eines monogonen LiDAR, die möglicherweise durch Nebel oder Staub verfälscht sind, zu validieren. Die Validierung der Daten findet ausschließlich im Sichtfeld des Lasers statt. Es werden keine Rückschlüsse auf den dreidimensionalen Aufbau der Hindernisse außerhalb der Lasersicht getroffen. Eine Minderung der Kollisionsgefahr mit überhängenden Hindernissen, die sich unter bzw. über der Scanebene befinden, kann nicht erreicht werden.

In dem von Garcia-Alegre et al. präsentierten Ansatz werden Features in einem monokularen Kamerabild mit korrespondierenden Messungen eines monogonen LiDAR kombiniert, um eine signifikante Aussage über mögliche Hindernisse zu erhalten. Das Resultat ist eine Detektion möglicher Hindernisformen im Kamerabild auf Basis von Schwellwertoperationen. Garcia-Alegre et al. kombinieren einen Laserdatensatz mit einem Bild, um eine Segmentierung des Bildes in Vorder- und Hintergrund zu erhalten. Der Fokus der Arbeit liegt auf der Gegenüberstellung von Segmentierungsansätzen. Eine Nutzung der gefundenen Segmente zur Hindernisdetektion wird nicht erläutert. Die Nutzung von nur einer Laserebene pro Bild lässt darauf schließen, dass die Hindernisoberfläche orthogonal zur Bodenebene angenommen wird, auf dem sich das Bodenfahrzeug bewegt.

4.2 Konzept

Für den Betrieb von unbemannten UA in einer unbekanntem Hinderniskulisse, wird, bedingt durch die Möglichkeit dreidimensional zu operieren, eine dreidimensionale Detektion von Hindernissen angestrebt. Begrenzt

durch die Größe von kleinen³ UA steht ausschließlich eine geringe Nutzlast für Sensoren der Hindernisdetektion zur Verfügung, was wiederum die Auswahl an dreidimensional messenden Sensoren, die ein Abbild der Umwelt bereitstellen, stark einschränkt.

Aktuell werden vor allem passive Stereokamerakonfigurationen und monogone LiDAR mit den erläuterten Nachteilen eingesetzt (siehe Abschnitt 4.1.1).

LiDAR ermöglichen im Allgemeinen eine Abbildung der Umwelt, welche geringer mit Fehlern belastet ist, als eine Abbildung, welche mit einer Stereokamera vergleichbarer Größe und Auflösung erstellt wurde. Darüber hinaus sind LiDAR durch die aktive Messmethode unabhängiger gegenüber externen Beleuchtungsänderungen. Dem gegenüber ermöglichen Kameras mit einem zweidimensionalen Sensorarray, im Vergleich zu weitverbreiteten LiDAR-Systemen mit einer bis wenigen Scanebenen [37, 83], im Allgemeinen eine Sicht auf die Umwelt, die vom Scanner nicht erreicht werden kann.

Erweiterungen der LiDAR mit mechanischen Komponenten ermöglichen ebenfalls eine Erweiterung des Sichtfeldes, weisen allerdings auch einen Zuwachs an benötigter Nutzlast und eine stark reduzierte Aktualisierungsrate auf.

Aufbauend auf dieser Situation sollen die Stärken der Sensoren Kamera (großes Sichtfeld und hohe Aktualisierungsrate bezüglich der Sichtfeldgröße) und LiDAR (hohe Messgenauigkeit bei hoher Einsatzbereitschaft bei unterschiedlichen Beleuchtungs- und Witterungssituationen) kombiniert werden, um die Schwächen Passivität und indirekte Entfernungbestimmung bei der Kamera und beschränktes Sichtfeld beim Laser zu kompensieren.

Um dies zu erreichen, sollen die Daten eines LiDAR mit den Bildern einer monokularen Kamera kombiniert werden, um eine dreidimensionale Sicht zu erzeugen, die über das Sichtfeld des LiDAR hinausgeht. Voraussetzung für dieses Ziel ist eine **kooperative Fusion** der beiden Datensätze für ein gemeinsames Sichtfeld, um die dort getroffenen Aussagen auf Bereiche außerhalb der Lasersicht zu übertragen. Das Resultat sind dreidimensionale Approximationen der Hinderniskulisse außerhalb der Lasersicht.

Das LiDAR dient dieser Konfiguration als Grundlage, da davon ausgegangen wird, dass dessen Einsatzbereitschaft bei dynamischen Umweltbedingungen höher ist als die einer Kamera. Die Kameradaten werden ausschließlich additiv der Fusion hinzugefügt, um auch bei widrigen Beleuchtungsverhältnissen eine rudimentäre Umweltwahrnehmung bereitzustellen.

³Nutzlast die für die Sensoren der Umweltwahrnehmung bereitsteht ist kleiner einem Kilogramm.

len, welche allein auf Laserdaten beruht.

Als Quelle für die Laserdaten wird eine monogone Konfiguration ausgewählt. Diese Konfiguration ist mechanisch recht einfach und ist in einigen, kleinen „commercial off-the-shelf“ Produkten zu finden. Der Einsatz von kleinen monogonen LiDAR-Systemen ist in kleinen UA direkt möglich [35, 65]. Die monogone Konfiguration ist mitunter eine der einfachsten Umsetzungen eines LiDAR und soll damit die Generalisierbarkeit des Ansatzes verdeutlichen. Ein Einsatz eines mehrzeiligen Systems ist ebenfalls durchaus möglich, soll aber nicht Inhalt dieser Arbeit sein.

Als Quelle für die Kameradaten wurde eine monokulare Kamera ausgewählt. Die Wahl, eine monokulare Kamera zu verwenden, hat zwei Gründe. Einerseits bringt ein Stereokamerasystem zusätzliches Gewicht mit in das Sensorsetup, wobei der möglicherweise erreichte Zusatznutzen durch die indirekte Entfernungsbestimmung fraglich ist [46]. Auf der anderen Seite führt der Einsatz einer monokularen Kamera ebenfalls zu einer Generalisierung des Ansatzes, was den Einsatz auf vielen Systemen erleichtert. Eine Erweiterung des Ansatzes mittels Stereokameradaten ist durchaus denkbar, soll in dieser Arbeit aber nicht betrachtet werden.

Die monokulare Kamera liefert skalierungsinvariante Abbildungen \mathbf{B} der Umwelt. Innerhalb der Bildebene können ohne weitere absolute Informationen nur relative Größenabschätzungen getroffen werden. Das monogone LiDAR erzeugt pro Scan \mathbf{A} eine einzelne Scanebene, in der absolute Entfernungsmessungen vorgenommen werden. Über den Zustand der Umwelt über- und unterhalb der Ebene können keine Aussagen getroffen werden. Sowohl monokulare Kamera als auch monogone LiDAR stellen mit einem Pixel $\mathbf{q}_i(x, y)$ bzw. einer Kreiskoordinate $\mathbf{p}_l(r, \Theta)$ ausschließlich zweidimensionale Daten bereit. Beide Sensoren können simultan von einem kleinen UA transportiert werden. Allein stellt aber keiner der Sensoren eine hinreichende dreidimensionale Sicht zur Verfügung.

Zur Einführung des angestrebten Konzepts wird von einem vereinfachten Aufbau ausgegangen. Ausgangspunkt ist, dass die Bildebene \mathbf{E}_B der Kamera und die Messungen des LiDAR in ein gemeinsames Koordinatensystem transformiert werden, im aktuellen Fall in das Koordinatensystem c der Kamera. Auf der Bildebene sind die relativen Ausdehnungen eines betrachteten Objekts sichtbar. Der Scan \mathbf{A} des LiDAR enthält für das vereinfachte Beispiel eine einzige Messung \mathbf{p}_l . Die Projektion \mathbf{q}_l von \mathbf{p}_l wird auf \mathbf{E}_B durch die sichtbaren Objektgrenzen des betrachteten Objekts umrahmt. Mit einer Lasermessung ist die Entfernung zu dem Objekt ermittelbar. Unter der Annahme, dass die Objekt Oberfläche des betrachteten Objekts einer Ebene \mathbf{E}_O entspricht und parallel zur Bildebene \mathbf{E}_B verläuft, lässt sich auf Basis von \mathbf{p}_l eine prädierte Ebenenapproximation \mathbf{E}'_O des betrachteten

Objekts erstellen. Eine Projektion der Objektumrisse von E_B auf E_O erlaubt eine Begrenzung der Oberflächenabmessungen. Dieses Vorgehen basiert auf dem Strahlensatz und ist in Abbildung 4.1 verdeutlicht. Abbildung 4.1

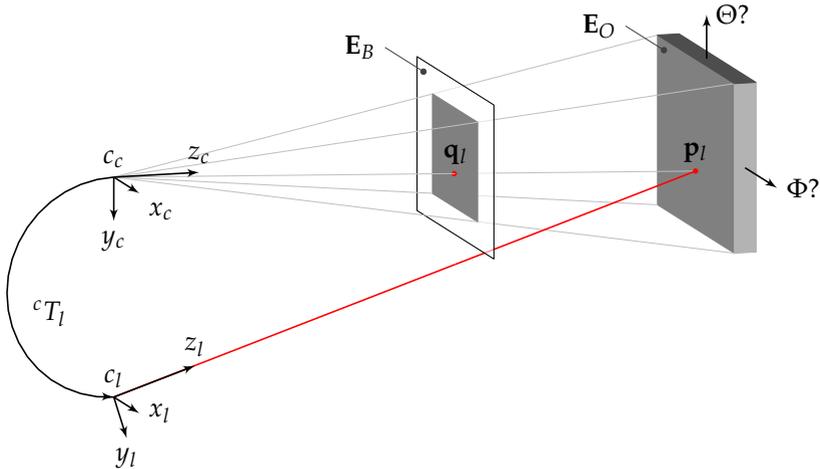


Abbildung 4.1: Approximation einer Objektoberflächenebene mittels einer Lasermessung und eines Kamerabildes.

zeigt, dass auf Grundlage einer Lasermessung und einem monokularen Kamerabild in Kombination mit getroffenen Annahmen, approximative Aussagen über Bereiche der Umwelt gemacht werden können auch wenn für diese Bereiche keine Lasermessungen, sondern ausschließlich skalierungs-invariante Kameradaten vorliegen. Es ist aber ebenso ersichtlich, dass eine Bestimmung der Objektoberflächenapproximation E'_O nur unter den Annahmen möglich ist, dass die betrachtete Objektoberfläche eine Ebene E_O ist und darüber hinaus parallel zur Bildebene verläuft. Im Hinblick auf den Einsatz des Konzeptes auf einem unbemannten Luftfahrzeug und der generellen Form von Objekten in der Umwelt ist die Verwendung dieser beiden Annahmen fragwürdig. Die getroffenen Annahmen werden im Folgenden näher betrachtet, um das anfänglich rudimentäre Modell des Konzeptes zu erweitern und eine generelle Lösung für eine Sichtfelderweiterung zu erzielen.

Bildebene verläuft parallel zur Objektoberflächenebene Unter Beibehaltung der Annahme, dass die Objektoberfläche eine approximative Ebene E'_O ist, ist auf Basis einer einzigen Lasermessung p_l nur ein Objekt-

punkt \mathbf{p}_l auf \mathbf{E}_O bekannt. Der Lage von \mathbf{E}'_O um \mathbf{p}_l bzw. die Winkel Φ und Θ ist unbekannt. Eine Herleitung der Lage ist ausschließlich über die Nutzung mehrerer unabhängiger Lasermessungen \mathbf{p}_l möglich. Für eine eindeutige Bestimmung der Lage von \mathbf{E}'_O sind mindestens drei linear unabhängige Messungen \mathbf{p}_l , wie in Abbildung 4.2, notwendig. Jede einzelne

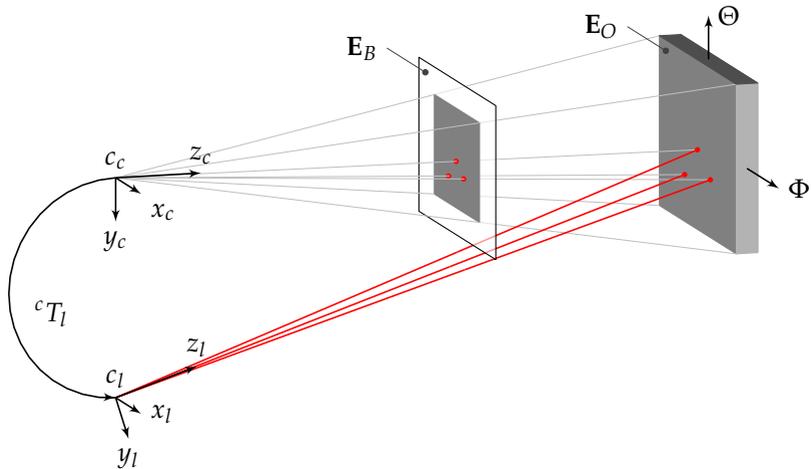


Abbildung 4.2: Idealsituation für die Sichtfelderweiterung mittels dreier Lasermessungen.

Lasermessung \mathbf{p}_l , die mittels eines monogenen LiDAR aufgezeichnet wurde, ist Bestandteil einer Punktwolke $\mathbf{p}_l \in \mathbf{C}_A$. Die Punktwolke \mathbf{C}_A enthält die Messungen einer Laserscanebene \mathbf{E}_A . Basiert \mathbf{C}_A auf einer \mathbf{E}_A und ist die Anzahl der Messungen \mathbf{p}_l in \mathbf{C}_A größer eins, lässt sich eine Schnittgerade \mathbf{g}_{OA} zwischen \mathbf{E}_O und \mathbf{E}_A ermitteln. Die Schnittgerade \mathbf{g}_{OA} ermöglicht die Bestimmung der Lage der Objektebene um eine Achse, die orthogonal zur Schnittgeraden verläuft und diese schneidet. In Abbildung 4.3 wird dies am Beispiel des Lagewinkels Θ verdeutlicht. Die Lage von \mathbf{E}'_O um Θ ist durch die Nutzung einer einzelnen \mathbf{E}_A möglich. Die Lage von \mathbf{E}'_O um Φ ist weiter unbekannt. Unter der Annahme, dass die Objektoberfläche \mathbf{E}_O eine Ebene ist, und der Schnitt zwischen Scanebene \mathbf{E}_A und \mathbf{E}_O folglich einer Gerade \mathbf{g}_{OA} ist, sind die Messungen bzw. Objektpunkte \mathbf{C}_A , die Grundlage von \mathbf{g}_{OA} sind, kollinear. Eine eindeutige Bestimmung der Lage für \mathbf{E}'_O auf Basis einer einzelnen \mathbf{E}_A ist somit nicht möglich.

Linear unabhängige Lasermessungen einer Objektoberfläche \mathbf{E}_O können durch das mehrmalige Abtasten dieser aus unterschiedlichen Positionen

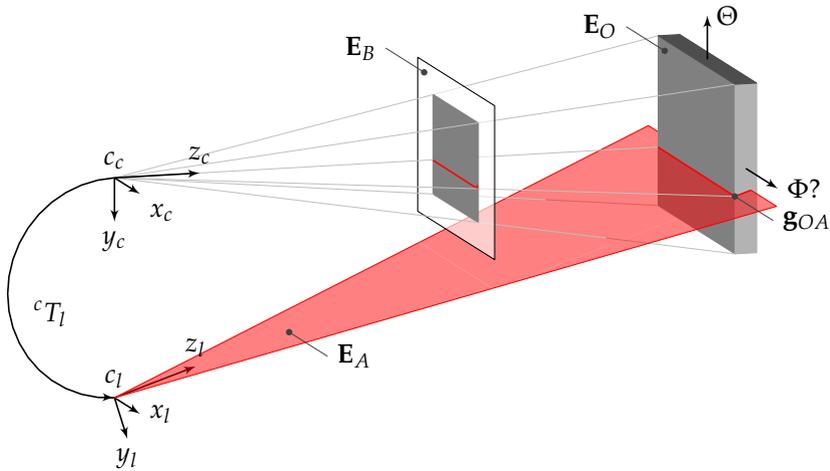


Abbildung 4.3: Approximation einer Objektoberflächenebene mit n -kollinearen Lasermessungen und einem Kamerabild.

und Orientierungen generiert werden. Um die Positions- und Lageänderungen des LiDAR zu erreichen, sollen keine zusätzlichen mechanischen Komponenten oder explizit für die Hindernisdetektion generierte Flugmanöver genutzt werden. Allein die Ausgleichsbewegungen des Luftfahrzeugs beim Beschleunigen, Abbremsen oder Manövrieren, sollen Basis für die benötigten Posenänderungen gegenüber dem Objekt sein.

Auf Grundlage mehrerer n zeitlich aufeinander folgender Scans $E_{A,t-n}$ bis $E_{A,t}$, die mit einem einzigen Bild der monokularen Kamera kombiniert werden, ist es möglich, ausgewählten Segmenten (potenzielle Objektoberflächen) auf der Bildebene jeweils eine Punktwolke $C_{A,t-n:t}$ von Lasermessungen zuzuordnen, deren Bestandteile jeweils linear unabhängig sind (siehe Abbildung 4.4). Auf Basis der jeweiligen Punktwolke wird eine Objektoberfläche E'_O pro Segment bestimmt, deren Lage und Position damit bekannt ist.

Die Objektoberfläche ist eben Diese Annahme trifft nur auf eine begrenzte Anzahl von Objekten bzw. Hindernissen zu, wird aber zur Vereinfachung für das folgende Fusionskonzept beibehalten. Im Folgenden wird erläutert, warum diese Annahme nur einen geringen Einfluss auf die resultierenden Ergebnisse der Hindernisdetektion hat und deshalb keine Einschränkung darstellt.

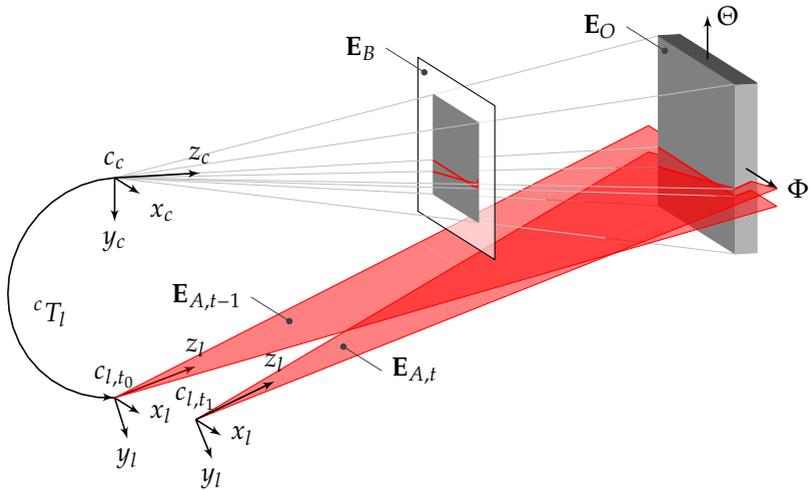
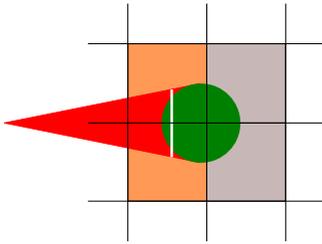


Abbildung 4.4: Approximation einer Objektoberflächenebene mit n-linear unabhängigen Lasermessungen und einem Kamerabild.

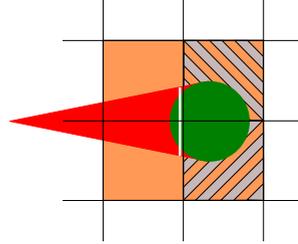
Das Ziel der hier präsentierten Hindernisdetektion ist die Erstellung einer dreidimensionalen Karte, welche den durch Hindernisse belegten Raum repräsentiert. Die eigentliche Form des Hindernisses ist dabei irrelevant. Eine häufig genutzte Umsetzung einer Hindernisrepräsentation ist eine metrisch diskretisierte Karte [7, 44]. In einer diskreten Karte, ob zwei- oder dreidimensional, repräsentieren die einzelnen Elemente Teile des Raumes, die jeweils belegt oder frei sind. Die Oberflächenstruktur eines Objektes, das ein Kartenelement belegt, ist dabei nur dann von Interesse, wenn die Strukturen die Auflösung der Karte überschreiten.

Abbildung 4.5 zeigt Beispiele, bei denen ein grünes zylindrisches Objekt, welches bspw. dem Querschnitt eines Baumes entspricht, mittels eines LiDAR (roter Kegel) vermessen wird. Die Abbildung zeigt die Draufsicht. Auf Basis der Messpunkte wurde eine Regressionsebene E'_O bestimmt, die in der Draufsicht weiß visualisiert ist. Die Regressionsebene E'_O wird als Approximation der Objektoberflächenebene E_O betrachtet. Die Überführung von E_O in das metrische, diskrete Belegtheitsgitter führt zu der Kennzeichnung des durch die Kartenelemente beschriebenen Raums entweder als belegt (orange) bzw. unbekannt/frei (grau).

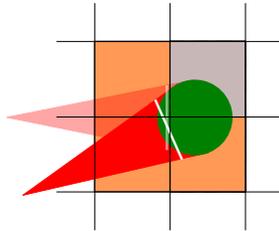
Abbildung 4.5a zeigt eine Situation, in der alle Rohmessungen des LiDAR in den Zellen liegen, welche ebenfalls von der Regressionsebene geschnitten werden. In diesem Fall führen sowohl Rohmessungen C_A als auch Re-



(a) Rohdaten und Regressionsebene führen zu identischer Belegung.



(b) Rohdaten und Regressionsebene führen zu abweichender Belegung.



(c) Fusion von mehreren Datensätzen führt zu einer Annäherung der Belegung.

Abbildung 4.5: Differenzierte Kennzeichnung eines Belegtheitsgitters ausgehend von Rohdaten oder einer Regressionsebene.

gressionsebene E'_O zu einer gleichen Belegung des Gitters. Abbildung 4.5b zeigt die Situation, in der sich Teile der Laserrohmessungen C_A in den Zellen befinden, die nicht von E'_O berührt werden. In diesem Fall werden die grau/orange gestreiften Zellen durch die Rohdaten C_A als belegt gekennzeichnet, während die E_R den Zustand (unbekannt) dieser Zellen nicht beeinflussen. Die Verwendung einer vereinfachten Oberfläche, in diesem Fall einer Ebene, führt nicht gezwungenermaßen zu einer identischen Belegung in der Hinderniskarte und damit nicht zu einer identischen Hindernisrepräsentation. Das vorgestellte Beispiel aus Abbildung 4.5b basiert auf einer Momentaufnahme der Situation. Metrische Belegtheitsgitter werden oft genutzt, um konkurrierende Sensordaten auf einer probabilistischen Basis zu fusionieren. Abbildung 4.5c zeigt eine Situation, die zeitlich nach der Situation in Abbildung 4.5b liegt. Durch eine Veränderung von Sensorposition oder -lage zum Hindernis ergibt sich eine neue Sicht auf das Hindernisszenario. Diese Veränderung hat zur Folge, dass Zellen, welche in Abbildung 4.5b durch die Daten der Regressionsgeraden nicht berührt wurden, zum Zeitpunkt von Abbildung 4.5c ebenfalls belegt werden.

Neben der additiven Belegung der Hindernisrepräsentation erlaubt die konkurrierende Kombination innerhalb eines probabilistischen Gitters auch die Freigabe von belegten Zellen. Die Abbildung 4.5c verdeutlicht eine Situation, in der die Regressionsebene ausschließlich Zellen schneidet, welche ebenfalls mindestens eine Rohmessung des Lasers aufweisen würden. Dieser Zusammenhang kann aber nicht garantiert werden. Ebenso kann eintreten, dass die Regressionsebene in Zellen liegt, welche keine Rohdaten enthalten. Es erfolgt eine Belegung von Zellen, welche tatsächlich frei sind. Dies tritt besonders dann auf, wenn die Ausdehnung des betrachteten Objekts über der Auflösung des Gitters liegt, die Oberflächenform von der Ebene abweicht und in den Bilddaten keine formgebenden Kanten detektiert werden können.

Eine vorsorgliche, konservative Belegung von freien Zellen ist im Allgemeinen sicherer als die Freigabe von vermutlich belegten Zellen. Allerdings führt eine übermäßige Belegung von Zellen auch zu einer Reduktion des möglichen Raumes, der zur Operation des UA genutzt werden kann. Dies führt besonders in engen Passagen zum Versperren möglicher Pfade. Aus diesem Grund wird angestrebt, mittels des probabilistischen Belegtheitsgitters die wahre Form der betrachteten Hindernisse zu rekonstruieren, um einerseits alle notwendigen Zellen zu belegen, andererseits das Ausmaß der Belegung möglichst auf die notwendigen Zellen zu beschränken.

Auf Grund des präsentierten Beispiels wird davon ausgegangen, dass es bei der Integration einer Momentaufnahme der **kooperativen Fusion** ge-

genüber der Integration von Rohdaten zu Unterschieden bei der Belegung des Belegtheitsgitters kommen kann. Ursache ist die Verwendung einer Approximation, die von den wahren Oberflächen abweichen kann. Eine aufsetzende **konkurrierende Fusion** mehrerer Momentaufnahmen, aus unterschiedlichen Sensorpositionen oder -lagen, ermöglicht eine fortschreitende Angleichung der Belegung, was schlussendlich in einer im hohen Maße übereinstimmenden Belegung des Gitters resultiert, unabhängig von der Art der Daten (Rohdaten oder Approximation).

Die konkurrierende Fusion mit einer probabilistischen metrischen Hindernisrepräsentation erlaubt neben der zeitlichen Kombination einer Datenquelle auch die simultane Kombination mehrerer Datenquellen. Im aktuellen Fall wird zusätzlich zu den Daten der Sichtfelderweiterung auch die gesamte Datenbasis des LiDAR integriert. Dies ist auf drei Gründe zurückzuführen. Erstens werden bei einer Sichtfelderweiterung nicht alle Laserdaten verarbeitet, sondern nur die Daten, welche im Blickfeld der Kamera liegen. Das Verwerfen von vorhandenen Messungen ist nicht zweckmäßig, weshalb die Laserdaten außerhalb der Kamerasicht ebenfalls für die Hindernisrepräsentation integriert werden sollten. Zweitens enthalten die Resultate für die Sichtfelderweiterung ausschließlich Approximationen. Um mögliche angesprochene Abweichungen zwischen Approximationen und Rohdaten für die Hindernisrepräsentation weiter zu berücksichtigen, werden neben den besprochenen Ansätzen die Laserrohdaten, welche innerhalb der Kamerasicht liegen, ebenfalls in der Hindernisrepräsentation beachtet. Der dritte Punkt für den Bypass der LiDAR-Daten ist die Bereitstellung einer rudimentären Hindernisrepräsentation auch bei einem Totalausfall der Kamera allein auf Basis des LiDAR. Die rudimentäre Hindernisrepräsentation hätte zwar nicht mehr den Umfang, welcher mit der angestrebten Sichtfelderweiterung erreicht würde, kann möglicherweise aber den Verlust des Gesamtsystems in einer Notfallsituation abwenden.

Das aktuelle Kapitel ist im Folgenden in drei Abschnitte unterteilt und orientiert sich an dem modularen Aufbau der gesamten Sichtfelderweiterung, welcher in Abbildung 4.6 skizziert ist. Zuerst wird auf die Vorverarbeitung der jeweiligen Rohdaten eingegangen. Darauf aufbauend erfolgt die Kombination der aufbereiteten Laser- und Kameradaten in der kooperativen Fusion. Die Resultate dieser Fusion bilden den Input für die konkurrierende Fusion, deren Ergebnis eine probabilistische metrische Hindernisrepräsentation ist, welche für die Pfadplanung eingesetzt werden kann.

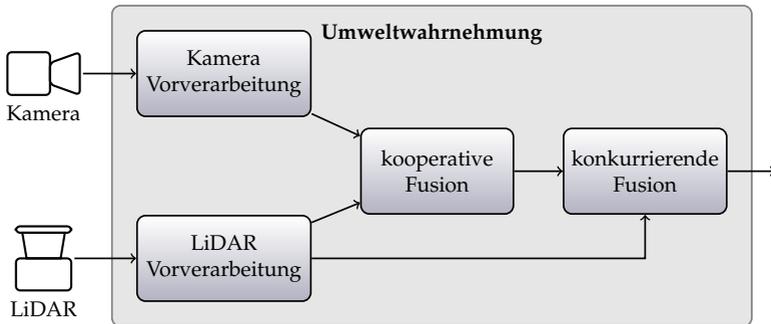


Abbildung 4.6: Blockdiagramm der Sichtfelderweiterung.

4.3 Vorverarbeitung

Die Vorverarbeitung der jeweiligen Sensordaten von Kamera und LiDAR ist auf die Anforderungen der kooperativen Fusion abgestimmt.

4.3.1 Vorverarbeitung des Kamerabildes

Die Vorverarbeitung des Kamerabildes hat in erster Linie das Ziel, ein Bild bereitzustellen, welches optimal zur Segmentierung geeignet ist. Dieser Anspruch basiert auf der Annahme, dass Objekte bzw. deren Oberflächen aus dem Objektraum als zusammenhängende Gruppen von Pixeln auf der Bildebene abgebildet sind. Ausgehend von einer mehr oder minder homogenen Oberflächengestaltung der Objekte sind die Intensitäten der einzelnen Gruppen ebenfalls homogen. Objekte, die eine geringe Entfernung zum Kamerazentrum besitzen, verdecken entferntere Objekte sowie den Hintergrund teilweise oder vollständig. Weiterhin wird angenommen, dass die Übergänge zwischen Objekten oder Objekten und dem Hintergrund zu einer Diskontinuität der Intensität führen. Ziel der Segmentierung ist es, Diskontinuität zu identifizieren, um einzelne Segmente S_i zu separieren, da davon ausgegangen wird, dass die einzelnen S_i , eine oder zu mindestens Teile einer Objektoberfläche repräsentieren.

Welches der in Abschnitt 2.4 aufgezeigten Verfahren verwendet wird, ist weniger relevant. Allein das erwartete Ergebnis der Segmentierung, zusammenhängende Regionen auf der Bildebene, für welche korrespondierende Lasermessungen vorliegen, ist von Interesse. Es bieten sich pixel-, region- als auch kantenbasierte Verfahren an. Modellbasierte Verfahren erscheinen in unbekanntem Gebieten weniger nützlich, da kein a-priori-Wissen über die geometrische Form der betrachteten Objekte vorliegt.

Eine ausschließliche Segmentierung einzelner Objektoberflächen stellt besonders bei Aufnahmen außerhalb von Gebäuden ein größeres Problem dar. Durch dynamische, nicht beeinflussbare Beleuchtungsänderungen und Schattenwurf können Intensitätsunterschiede zwischen einzelnen Objektoberflächen sehr gering ausfallen oder gar verschwimmen, was eine Separation erschwert bzw. verhindert. Diese Problematik betrifft alle Segmentierungsansätze, die auf Intensitätsunterschieden basieren.

Garcia Alegre [31] verwendet eine regionbasierte Segmentierung, welche als Keimpunkte, die auf die Bildebene projizierten Lasermessungen verwendet. Als Abbruchkriterium nutzt sie eine feste Anzahl von Iterationen des Algorithmus bzw. eine feste im Vorfeld bestimmte Anzahl von erwarteten Segmenten. Die Verwendung von Abbruchkriterien wie die Anzahl der Iterationen oder möglicher Segmente ist nur bedingt für den Einsatz in unbekanntem, dynamisch beleuchteten Umgebungen umsetzbar, weil die wahre Anzahl der benötigten Iterationen bzw. Segmente unbekannt ist

In der hier präsentierten Arbeit, wird eine kantenbasierte Segmentierung (siehe Abschnitt 2.4.2) eingesetzt, da für den verwendeten Canny-Algorithmus kein Abbruchkriterium benötigt wird und die benötigte Parametrisierung automatisiert bestimmt werden kann. Diese macht den Einsatz für Bilddaten, welche starken Dynamikschwankungen unterliegen, etwa außerhalb von Gebäuden, interessant und ermöglicht eine automatisierte Detektion der relevanten Kanten.

Eine alleinige Nutzung eines Kantendetektors, wie Canny (siehe Abschnitt 2.4.2), direkt auf den Rohbildern ist dennoch nicht zu empfehlen. Einerseits nutzt die automatisierte Parametrisierung des Canny nur die gegebene Dynamik des Bildes, aber nicht den maximal möglichen Dynamikumfang von diesem, was möglicherweise relevante Kanten verschleiert. Auf der anderen Seite handelt es sich bei einem Kantendetektor um einen Pixel- und lokalen Operator, welcher zur Identifikation der gesuchten Intensitätsdiskontinuitäten jedes Pixel des Bildes mindestens einmal betrachten muss. Dieses führt zu einem erhöhten Prozessieraufwand [49, S. 143]. Eine Reduktion des Aufwands mittels einer Datenkompression bzw. einer Reduktion der betrachteten Pixelanzahl ist anzustreben.

Um die beiden angesprochenen Punkte zu erreichen, ist der Kantendetektor Teil einer Filterkette, die in Abbildung 4.7 skizziert ist. Die Identifizierung der Kanten erfolgt mittels lokaler Intensitätsdiskontinuitäten. Je prägnanter die Diskontinuitäten ausfallen, desto konkreter lässt sich eine Kante identifizieren. Die Intensitätsdiskontinuität zwischen dem hellsten Pixel und dem dunkelsten Pixel eines Bildes wird als Kontrastumfang oder Dynamik bezeichnet. Wird bei einer Aufnahme nicht die gesamte Dynamik genutzt, die theoretisch zur Verfügung steht, kann durch eine homo-

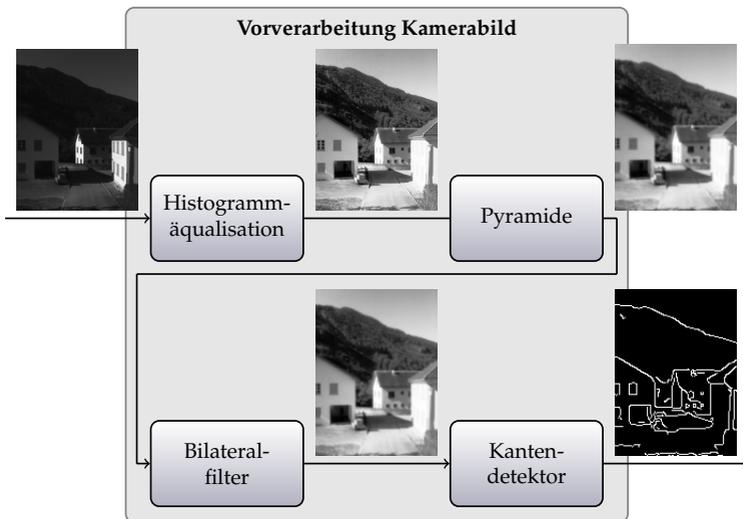


Abbildung 4.7: Blockdiagramm der Vorverarbeitung der Kamerabilder.

gene Transformation der jeweiligen Intensitäten eine Steigerung der Dynamik erreicht werden. Ansätze zur Steigerung des Kontrastumfangs von Bildern, mit nicht vollständiger Ausnutzung der Dynamik, sind die Histogrammspreizung und die Histogrammäqualisation [49, S. 265], [41]. Das erste Element der Filterkette aus Abbildung 4.7 ist eine Histogrammäqualisation mittels deren Hilfe der geringe Dynamikumfang des Ausgangsbilds auf das maximale Maß gesteigert wird.

Wie erwähnt, benötigt die Betrachtung jedes Pixels bei einem Pixel- bzw. lokalen Operator wie einem Kantendetektor einen gewissen Aufwand, den es zu reduzieren gilt [49, S. 143]. Ein gängiges Verfahren zur Herabsetzung der Bildauflösung ist die Gauß-Pyramide. Das Ausgangsbild bildet die nullte Ebene. Bei jeder Erhöhung der Pyramidenebene wird die Auflösung gegenüber der darunter liegenden Ebene halbiert. Dazu wird der Mittelwert von vier benachbarten Pixeln bestimmt, um den Wert des korrespondierenden Pixels auf der darüber liegenden Pyramidenebene zu ermitteln [49, S. 456]. Je nach Anzahl der verwendeten Ebene dezimiert sich die Anzahl der zu betrachtenden Pixel für die Folgeschritte in der Filterkette. Die Darstellungen in Abbildung 4.7 weisen aus Gründen der Visualisierung eine konstante Größe auf. Die Auflösung der Ergebnisbilder nach der Gauß-Pyramide ist allerdings halbiert.

Die Histogrammäqualisation führt durch die Steigerung des Dynamikum-

fangs, nicht nur zur Sichtbarkeit von relevanten Kanten, sondern ebenfalls zu erhöhtem Rauschen in homogenen Flächen, was nachteilig für die Kantendetektion ist. Durch den Einsatz der Gauß-Pyramide wird durch die Mittlung der benachbarten Pixel zusätzlich eine Glättung erreicht, die einem Tiefpassfilter gleichkommt und zu einer Reduktion des Rauschens führt.

Auf der einen Seite kann die wiederholte Anwendung der Pyramide durch die Glättung zu einem Verlust von relevanten Kanten führen. Aus diesem Grund wird in der hier dargelegten Filterkette eine Pyramide mit nur zwei Ebenen verwendet. Dies ermöglicht das Vierteln der zu betrachtenden Pixelanzahl. Auf der anderen Seite genügt diese Anwendung nicht, um das Rauschen hinreichend in den homogenen Flächen einzudämmen.

Um einerseits die Kanten zu erhalten und andererseits eine weitere Glättung der homogenen Flächen zu erhalten, wird als dritter Schritt der Filterkette ein Bilateralfilter [91] eingesetzt.

Als letzter Schritt der Filterkette erfolgt der Einsatz eines Kantendetektors, in dem hier dargelegten Fall wird ein Canny-Algorithmus verwendet (siehe Abschnitt 2.4.2), der mit einer Histogramm-Konkavitätsanalyse (siehe Abschnitt 2.4.2) automatisch parametrisiert wird.

Das Resultat der Filterkette bzw. der Vorverarbeitung der Kamerabilder ist ein binäres Bild, in dem die entsprechenden Kantenpixel markiert sind. Dieses Bild bildet das Eingangssignal für die kooperative Fusion im Speziellen für die Segmentierung.

4.3.2 Vorverarbeitung der Laserdaten

Die Vorverarbeitung der Laserdaten berücksichtigt zwei Punkte:

- Kamera und Laserdaten müssen in ein gemeinsames Koordinatensystem überführt werden, wobei die Asynchronität der Datenaufzeichnung zwischen Kamera und Laser sowie die serielle Datengenerierung des LiDAR beachtet werden muss
- die Messungen eines monogenen Scans \mathbf{A} liegen in einer Ebene E_A und enthalten für ein Objekt mit planarer Oberfläche ausschließlich kollineare Daten. Zur Herleitung von linear unabhängigen Messungen ist die Kombination mehrerer \mathbf{A}_t , welche nicht zusammenfallen, notwendig.

Für die Sichtfelderweiterung wird das Kamerakoordinatensystem als Basissystem verwendet. Die Überführung der Messungen aus dem Laserkoordinatensystem in Kamerakoordinaten wird mit der relativen Transfor-

mation ${}^c\mathbf{T}_l$ erreicht, welche in Abschnitt 3.1 hergeleitet wurde. Eine direkte Anwendung von ${}^c\mathbf{T}_l$ kann allerdings nur unter der Annahme erfolgen, dass zwischen den Messungen eines Laserscans \mathbf{A} und der Aufnahme der Kamera keine Bewegungen des Sensoraufbaus gegenüber der Umwelt bzw. umgekehrt erfolgen. Diese Annahme kann für einen seriell arbeitenden LiDAR und einen asynchronen Betrieb mehrerer Sensoren, welche an einem dynamischen UA montiert sind, im Allgemeinen nicht aufrechterhalten werden.

Die Messungen des LiDAR werden, wie in Abschnitt 2.3.2 erläutert, innerhalb eines Laserscans \mathbf{A} nicht simultan erzeugt, so dass während des einmaligen Abtastens des Sichtfeldes eine Zeitspanne Δt_A benötigt wird. Die Aufnahmen der Kamera und des Lasers sind u.a. aufgrund unterschiedlicher Aktualisierungsraten asynchron, so dass zwischen den jeweiligen Aufnahmen ebenfalls eine Zeitspanne Δt_f liegt. Innerhalb der jeweiligen Zeitspannen finden Bewegungen des UA gegenüber der Umwelt statt. Eine direkte Anwendung von ${}^c\mathbf{T}_l$ führt zu einer Vernachlässigung der Bewegungen während Δt_A sowie Δt_f und führt damit zu einer Verfälschung der aufgezeichneten Sensordaten.

Auf Basis der Gleichungen 2.13, 2.11 und 2.20 werden die einzelnen Lasermessungen zum Zeitpunkt ihrer Erzeugung t_m separat in das lokale Weltkoordinatensystem o transformiert.

$$\mathbf{p}_o = ({}^o\mathbf{T}_{l,t_m})^{-1} \mathbf{p}_l, \quad (4.1)$$

Dort werden sie gesammelt und als Punktwolke \mathbf{C}_A zum Zeitpunkt der Kameraaufnahme t_c mit den dann aktuellen Navigationsdaten (siehe Abschnitt 2.13) in das Kamerasystem und schlussendlich per Projektion (Gleichung: 2.31) auf die Bildebene überführt.

$$\mathbf{p}_c = {}^o\mathbf{T}_{c,t_c} \mathbf{p}_o, \quad (4.2)$$

Diese Prozedur ist im Blockdiagramm in Abbildung 4.8 verdeutlicht, wobei die Aufbewahrung der Laserdaten im lokalen Weltkoordinatensystem durch einen FIFO⁴ Puffer und der Abruf durch einen Trigger, welcher zum Zeitpunkt t_c ausgelöst wird, visualisiert ist. Durch die Transformation der Laserdaten in das o -Koordinatensystem unter Verwendung der aktuell vorliegenden Lage- und Positionsdaten des UA werden die Bewegungen, die zwischen den einzelnen Messungen eines Scans oder in der Zeitspanne Δt_f zwischen einem Laserscan und einer Kameraaufnahme vom UA vollführt werden, berücksichtigt und es findet keine Verfälschung der Sensordaten

⁴First in first out Speicherkonzept. Daten werden in der Reihenfolge aus dem Speicher entfernt, wie sie in den Speicher verbracht wurden.

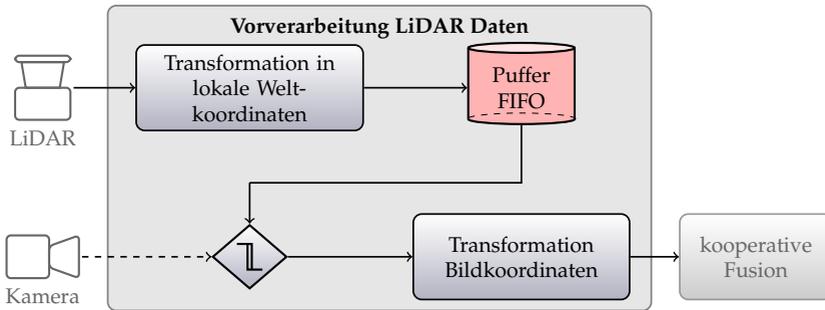


Abbildung 4.8: Blockdiagramm der Vorverarbeitung der Laserdaten.

statt. Voraussetzung für diesen Weg über das i -Koordinatensystem ist die Annahme, dass die Lage und Positionsdaten hinreichend genau sind, wovon in diesem Zusammenhang ausgegangen wird.

Die Hinterlegung der Laserdaten im o -Koordinatensystem ermöglicht nicht nur die Berücksichtigung der UA Bewegung für einen zuletzt erzeugten Scans, sondern ebenso für eine größere Anzahl n Scans \mathbf{A}_t . Eine Punktwolke $\mathbf{C}_{A,t-n:t}$ im Puffer, welche die Messungen mehrerer Scans enthält, wobei jede Messung separat in den Puffer überführt wurde, kann zum Zeitpunkt einer Bildaufnahme t_c ebenso auf die Bildebene überführt werden wie die Messungen eines einzelnen Scans. Dieses Vorgehen ermöglicht die Bereitstellung von linear unabhängigen Messungen, wenn davon ausgegangen wird, dass die aufgezeichneten Laserrohmessungen ebenfalls linear unabhängig sind, wovon hier ausgegangen wird.

Das Resultat der Filterkette bzw. der Vorverarbeitung der Laserdaten ist eine Punktwolke. Die Punktwolke enthält 1 bis n Scans deren Messungen in das Kamerakoordinatensystem transformiert wurden. Diese Punktwolke bildet das Eingangssignal für die kooperative Fusion.

4.4 Kooperative Fusion

Das Ziel der kooperativen Fusion von monokularen Kamerabildern und monogonen Laserscannerdaten ist die Bereitstellung von planaren Objektflächenapproximationen, welche Auskunft über Position und Lage von realen Objektflächen im Objektraum geben. Speziell von Interesse sind die Teile der Objektflächen, welche in den Teilen des Objektraumes liegen, für die keine Lasermessungen vorliegen.

Die kooperative Fusion lässt sich in zwei Arbeitsschritte zerlegen: Im ers-

ten Schritt wird die Zerlegung der Bildebene in Segmente sowie die Identifikation der Segmente, für welche Lasermessungen vorliegen, vorgenommen. Im zweiten Schritt erfolgt die Approximation der Segmente im Objektraum, was die Bestimmung der Position, Lage und Grenzen beinhaltet. Ein Blockdiagramm dieses Vorgehens zeigt Abbildung 4.9. Neben den aufgezeigten Arbeitsschritten enthält das Blockdiagramm einen weiteren Block „Anpassung der LiDAR-Auflösung“. Dieser Block dient zur weiteren Aufbereitung eines Teils der Laserscannermessungen, welcher ausschließlich zur Segmentierung benötigt wird und wird daher als Teil der Segmentierung betrachtet. Sowohl die Segmentierung als auch die Oberflächenapproximation benötigen Lasermessungen. Der Segmentierung wird ausschließlich der zuletzt erzeugte Scan, der letzte Datensatz (engl. tail) im FIFO, zugeführt und der Oberflächenapproximation eine unbestimmte Anzahl n , welcher maximal der Anzahl von Datensätzen im FIFO Puffer der Laser-Vorverarbeitung entspricht.

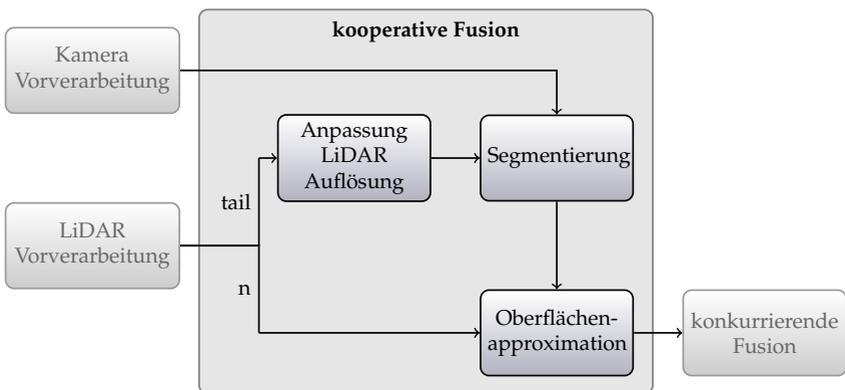


Abbildung 4.9: Blockdiagramm der kooperativen Fusion von Kamerabilddaten und Laserscannerdaten.

4.4.1 Segmentierung

Die Eingänge der Segmentierung bilden die vorverarbeiteten Datenströme von Kamera und der zuletzt erzeugte Scan des Laserscanners. Bei den Daten der Kamera handelt es sich um ein Ergebnisbild eines Canny-Kantendetektors, welches Diskontinuitäten der Intensitäten zwischen homogenen Bildbestandteilen hervorhebt. Der Canny-Kantendetektor gibt keine Kur-

ve oder Segmente zurück, sondern ein Bild, in dem lokale Diskontinuitäten der Intensität verstärkt und homogene Grauwertverläufe geglättet und binarisiert sind, so dass am Ende ausschließlich die Diskontinuitäten bzw. Kanten gekennzeichnet sind. Jedes Kantenpixel (Diskontinuität der Intensität) steht innerhalb des Bildes für sich allein. Ziel der Segmentierung ist, aus den einzelnen, zueinander beziehungslosen Pixeln zusammenhängende Konturen zu generieren, welche mögliche Segmente begrenzen, um einen gezielten Zugriff auf einzelne Segmente zu ermöglichen.

Die angestrebte kooperative Fusion der Sensordaten ist allerdings nur für Segmente möglich, für welche ebenso korrespondierende Lasermessungen vorliegen. Dieses trifft nicht auf alle verfügbaren Segmente eines Bildes zu. Eine Segmentierung der Bildbereiche, für die keine Lasermessungen vorliegen, ist nicht notwendig und sollte im Hinblick auf die Performanz des Ansatzes bewusst ausgeschlossen werden. Folglich wird angestrebt, ausschließlich die Kantenpixel im binären Kantenbild zu identifizieren, denen Lasermessungen zugeordnet werden können.

Die Einschränkung, die Segmentierung ausschließlich auf Bildbereiche vorzunehmen, für welche Lasermessungen vorliegen, erfordert eine Zuführung der Laserdaten in das Modul der Segmentierung. Die Laserdaten sollen als Saatkorn bzw. Ausgangspunkt für die Segmentierung verwendet werden, um Konturen zu definieren, welche interessante Segmente umrahmen.

Überblick

Segmentierung ist die analytische Aufteilung eines Bildes in intern homogene und extern heterogene Mengen sowie der konzentrierte bzw. differenzierte Zugriff auf eine oder mehrere dieser Mengen. Ein Segment wird, basierend auf der Kantendetektion, durch eine Kontur⁵ bestimmt, die es umgibt. Eine Zuordnung der Kantenpixel zu einer Kontur kann mit lokalen oder globalen Ansätzen erfolgen [33].

Lokale Ansätze nutzen wie die präsentierten Kantendetektoren lokale Operatoren, welche die engere Nachbarschaft (3×3 oder 5×5) jedes Pixels $\mathbf{q}(u, v)$ untersuchen, das als Kantenelement identifiziert wurde. Betrachtet werden die Kantenstärke $\nabla I(u, v)$ sowie Kantenrichtung $\Theta(u, v)$. Pixel mit vergleichbaren Merkmalen bilden überschneidungsfreie Mengen nach vorab definierten Kriterien. Auf Grundlage der jeweiligen Punktmenge wird eine Kante mit den dementsprechenden Eigenschaften bestimmt. Schlussendlich wird im idealen Fall ein Kantenschluss erreicht und die Kontur beim inertialen Pixel geschlossen. Die geschlossene Kontur um-

⁵Menge aller Umrisspunkte einer Fläche.

rahmt ein mögliches Segment und erlaubt damit die vollständige Segmentierung. Ein geschlossener Kantenzug ist der Idealfall, der nicht garantiert werden kann. Je nach Wahl der Kriterien ändert sich die Zusammensetzung einer Menge und die daraus resultierende Kontur. Größere Lücken innerhalb der Konturen führen zu einer nicht eindeutigen Abgrenzung von Segmenten.

Globale Ansätze bestimmen Konturen unter der Berücksichtigung des gesamten Bildes bzw. aller als Kante identifizierten Pixel. Ein Beispiel für einen globalen Ansatz ist die Hough-Transformation [41, S. 86], welche Kantenpixel zu komplexeren Strukturen wie Geraden, Kreisen oder anderen parametrisierbaren geometrischen Figuren kombiniert. Die Hough-Transformation ist ein modellbasierter Ansatz [49, S. 460]. Jede Bewertung, ob $\mathbf{q}(u, v)$ zu dieser oder jener geometrischen Figur gehört, erfordert eine separate Hough-Transformation mit den jeweils definierten Parametern. Es werden nicht beliebige Figuren, sondern ausschließlich im Vorfeld parametrisierte Figuren identifiziert. Eine Auswahl von erforderlichen Figuren setzt a-priori Wissen voraus.

Ideale Segmentierung

Ausgehend von der idealisierten Situation in Abbildung 4.3, in der ein einziges gut ausgeleuchtetes Objekt mit einer homogenen, planaren Oberfläche von Kamera und Laserscanner vor einem ebenfalls homogenen Hintergrund detektiert wird, ist die Kombination von einzelnen Kantenpixeln mit bspw. einem lokalen Ansatz rudimentär.

Die Laserebene $E_{A,t}$ des aktuellen Scans A_t schneidet die Oberfläche des im Sichtfeld befindlichen planaren Objekts. Die Abmessungen der geschnittenen planaren Objektoberfläche E_O innerhalb $E_{A,t}$ sind damit bekannt. Die absolute Oberflächenausdehnung ober- und unterhalb von $E_{A,t}$ kann auf dessen Basis nicht bestimmt werden.

Das Objekt wird ebenfalls auf der Bildebene der Kamera abgebildet. Mittels einer Kantendetektion werden die Objektoberfläche E_O sowie der Hintergrund als homogene Flächen geglättet und nur der Übergang zwischen beiden ist als Kante sichtbar. Die Kantenpixel umgeben das Objekt vollständig.

Die Schnittgerade g_{OA} zwischen E_A und der Objektoberfläche wird mit $({}^cT_l)^{-1}$ auf die Bildebene projiziert (siehe Abschnitt 4.3.2).

Ausgehend von dieser Situation, kann mit einer lokalen Konturverfolgung, die an einem der Schnittpunkte zwischen dem projizierten Laserschnitt und einem Kantenpixel initiiert wird, die Kontur ober- und unterhalb des Schnittes bestimmt werden. Ergebnis ist eine geschlossene Kontur, die das

Segment vollständig umrahmt und damit einen segmentbezogenen Zugriff ermöglicht.

Gegebene Situation der Bilddaten

Die angestrebte Segmentierung von Bildmaterial, welches in natürlichen Umgebungen außerhalb von Gebäuden aufgezeichnet wurde, weicht in seiner Qualität von den idealisierten Bedingungen ab. Eine reale Szenerie enthält im Allgemeinen mehrere Objekte, deren Oberflächen nur bedingt homogen sind. Eine nicht ideale Ausleuchtung, Oberflächenfarbe oder Schattenspiel sorgen für identische oder ähnliche Intensitäten bei unterschiedlichen Objekten im Bild.

Das Resultat sind Verschleifungen von Intensitätsdiskontinuitäten zwischen einzelnen Objektoberflächen, was wiederum die Identifikation von Kanten erschwert oder gar verhindert. Eine geschlossene Kontur bzw. eine Kontur, welche ausschließlich ein einzelnes Objekt umgibt, kann nicht garantiert werden. Vielmehr muss davon ausgegangen werden, dass eine Kontur mehrere Objekte umrahmt, welche sich in unterschiedlichen Entfernungen zur Bildebene befinden.

Auf der anderen Seite muss davon ausgegangen werden, dass eine abgebildete Objektoberfläche durch Beleuchtung und ihre Textur nicht allein durch ein Segment repräsentiert wird. Die Segmentierung basiert allein auf der Homogenität der jeweiligen Oberfläche. Eine Objektoberfläche kann durch mehrere Segmente dargestellt werden. Segment wird nicht als Synonym für eine einzelne Objektoberfläche verstanden.

Der Einfluss von Texturen und Beleuchtung auf die Intensitätsverläufe eines Bildes und damit auf die Anwesenheit von Kanten lässt keine Aussage darüber zu, ob es sich bei einer Kontur um eine räumliche Änderung der Objektform handelt oder ob es sich um eine Änderung der Textur auf einer planaren Oberfläche handelt. Basiert ein Segment bzw. dessen Kontur auf der Heterogenität einer Textur auf einer planaren Oberfläche, fällt das identifizierte Segment im Allgemeinen kleiner aus, als ein Segment, welches die gesamte Objektoberfläche ohne Textur enthalten würde.

Konturen, die auf Texturen basieren, führen zu einer Reduktion der Bereiche auf der Bildebene, für welche theoretisch korrespondierende Lasermessungen vorliegen. Eine Differenzierung, ob die Heterogenität der Intensität zwischen einzelnen Pixeln auf formgebenden oder texturalen Kanten basieren, ist nicht möglich. Aus konservativer Sicht wird jede identifizierte Kante als formgebende Objektkante behandelt, was zu einer Reduktion des Approximationsraums führt, dadurch aber auch Falschaussagen vorbeugt.

Konturbestimmung auf Basis von Laserprojektionen

Die Abweichungen zwischen dem idealen Zustand und der gegebenen Situation machen den Einsatz einer lokalen Konturverfolgung nicht empfehlenswert. Das iterative Identifizieren von Kantenpixeln basierend auf der Nachbarschaft von zuvor identifizierten Kantenpixeln führt bei größeren Lücken oder Verzweigungen zu Situationen, die von dem Algorithmus nur bedingt gelöst werden können. Es wird davon ausgegangen, dass das Ziel, geschlossene Konturen zu ermitteln, die jeweils nur eine einzige Objektoberfläche enthalten, nicht zufriedenstellend umgesetzt wird.

Um einerseits nur die Segmente und damit deren Konturen zu bestimmen, die korrespondierende Lasermessungen enthalten, und andererseits, die Gefahr zu reduzieren, dass ein Segment mehrere Objektoberflächen in unterschiedlichen Entfernungen enthält, wird kein iteratives Verfahren wie bei einem lokalen Ansatz verwendet. Vielmehr soll jedes Kantenpixel unmittelbar von einer projizierten Lasermessung abgeleitet werden.

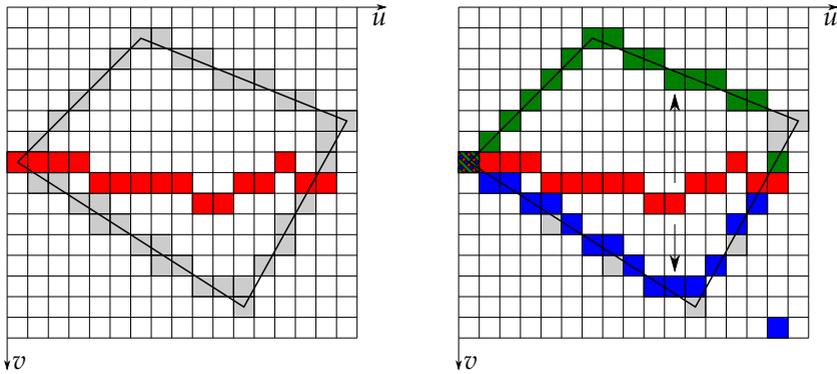
Eine Verwendung der Lasermessungen als Ausgangspunkt setzt eine Transformation dieser auf die Bildebene voraus, wie sie in der Vorverarbeitung der Laserscannerdaten (siehe Abschnitt 4.3.2) vorgestellt wurde. Die auf die Bildebene überführten Lasermessungen \mathbf{q}_l sind Element von \mathfrak{R} . Das Bild einer digitalen Abbildung ist ein diskretes Array. Für eine Berücksichtigung der projizierten Messungen innerhalb des Bildes ist eine Transformation von \mathbf{q}_l zu $\mathbf{q}_l(u, v)$ notwendig, wobei $u, v \in \mathfrak{N}$. Die Diskretisierung von \mathbf{q}_l geht mit einem Verlust an Genauigkeit einher.

Innerhalb der Laserebene $\mathbf{E}_{A,t}$ ist die Ausdehnung einer betrachteten Objektoberfläche bekannt. Ober- und unterhalb von $\mathbf{E}_{A,t}$ sind die Ausmaße des Objektes dagegen unbekannt. Bei der Projektion des Schnittes \mathbf{g}_{OA} auf die Bildebene (vergleichbar zu Abbildung 4.3), bietet sich folglich eine Suche nach möglichen Kantenpixeln orthogonal zur stärksten Ausdehnung von \mathbf{g}'_{OA} an. Eine Suche orthogonal zu \mathbf{g}'_{OA} verläuft nicht notwendigerweise parallel zur Orientierung einer der Bildachsen, was einen erhöhten Prozessierungsaufwand erfordert. Es wird daher eine Suche entlang der Bildachsen angestrebt. Entlang welcher Bildachse (u, v) die Suche vorangetrieben wird, ist durch die Lage von \mathbf{g}'_{OA} bezüglich der Bildachsen bedingt. Um dem angestrebten Ideal, einer Suche orthogonalen zu \mathbf{g}'_{OA} möglichst nahezukommen, wird zwischen \mathbf{g}'_{OA} und der Suchrichtung, bzw. der Bildachsen, ein möglichst großer Winkel α angestrebt, wobei $\alpha > 45^\circ$.

Abbildung 4.10a präsentiert ein Beispiel, in welchem die Orientierung eines projizierten Schnittes \mathbf{g}'_{OA} zur Achse u einen größeren Winkel α als zur Achse v aufweist. Folglich ist die angestrebte Suchrichtung parallel zu v . Die Raster auf der u -Achse werden im Weiteren auch als Spalten, die der

v-Achse als Zeilen bezeichnet.

Die Zuordnung von unabhängigen Kantenpixeln zu einer Kontur ausge-



(a) Ausgangssituation: Bildebene mit identifizierten Kantenpixeln und projizierten Lasermessungen.

(b) Suche von Kantenpixeln entlang der Spalten, für die Lasermessungen vorliegen.

Abbildung 4.10: Identifikation von Konturpixeln mittels korrespondierenden Lasermessungen.

hend von projizierten Lasermessungen wird anhand der Projektion eines homogen gefärbten Objektes auf die Bildebene (vergleichbar Abbildung 4.3) verdeutlicht. Abbildung 4.10a zeigt die Ausgangssituation auf der Bildebene. Vorhergehende Abläufe sind das Projizieren des Objektes auf die Bildebene, der Einsatz eines Kantendetektors, welcher die Intensitätsdiskontinuitäten (Kantenpixel) zwischen Objekt und Hintergrund hervorhebt sowie die Projektion der Lasermessungen auf die Bildebene. Die Kantenpixel sind grau dargestellt und umrahmen die zum besseren Verständnis kontinuierliche Kontur des abgebildeten Objekts. Die Pixel $\mathbf{q}_l(u, v)$, für die korrespondierende Lasermessungen vorliegen, sind rot markiert.

Ausgehend von jeder projizierten Lasermessung $\mathbf{q}_l(u, v)$ werden innerhalb der jeweiligen Spalte u zwei Kantenpixel bestimmt (siehe Abbildung 4.10b). Die beiden Pixel entsprechen den nächstliegenden Kantenpixeln ober- $\mathbf{q}_c(u, v - n)$ (grün) und unterhalb $\mathbf{q}_c(u, v + m)$ (blau) von $\mathbf{q}_l(u, v)$. Wird ausgehend von $\mathbf{q}_l(u, v)$ innerhalb von u entlang einer der Suchrichtungen kein Kantenpixel identifiziert wird für diese Suchrichtung der Bildrand als Kantenpixel definiert. Wurde das von $\mathbf{q}_l(u, v)$ (rot/blau/grün) belegte Pixel ebenfalls als Kantenelement identifiziert, wird die Suche unverzüglich beendet. Enthält eine Spalte keine projizierte Messung des Lasers,

wird keine Suche gestartet. Die identifizierten Kantenpixel ober- und unterhalb von A_t werden jeweils zu einer Menge U und D zusammengefasst, wobei $U \cap D = \emptyset$.

Segmentkonstruktion

Die Punktmenge U und D die ober- und unterhalb von A_t angeordnet sind, enthalten jeweils Konturteile mehrerer Segmente. Aussagen über einzelne Segmente können aus der jeweiligen Menge nicht entnommen werden. Der Aufbau einer Kontur pro Segment setzt eine Separation der Mengen U und D in Teilmengen voraus, wobei jede Teilmenge nur Konturelemente eines Segments aufweisen sollte.

Ein Segment beschreibt eine Fläche auf der Bildebene mit einer homogenen Intensität, von der in der präsentierten Arbeit angenommen wird, dass sie im Objektraum eine zusammenhängende planare Fläche abbildet. Die Grenzen des Segments bilden, in den Objektraum projiziert, ebenfalls die Grenzen der dortigen Fläche. Für die Transformation der Segmentgrenzen von der Bildebene in den Objektraum wird für ein Segment eine generelle Form angestrebt, die eine einfache Parametrisierung erlaubt.

Als eine einfache generalisierbare Parametrisierung eines Segments wird in der präsentierten Arbeit ein Trapez verwendet. Die Grundseiten⁶ des Trapezes verlaufen parallel zu der zuvor angewendeten Suchrichtung, daher entweder parallel zur Orientierung der Spalten oder der Zeilen. Die Schenkel eines Trapezes sind Geraden und werden jeweils durch zusammenhängende Teilmengen der Mengen U und D beschrieben. Für die Herleitung, welche Bestandteile der Mengen U und D zu welchem Segment bzw. Trapez gehören, ist eine jeweilige Zerlegung von U und D in Untermengen notwendig. Jede Untermenge repräsentiert eine Menge von Punkten, die einer approximierten zweidimensionalen Geraden auf der Bildebene zugeordnet ist.

Die Aufteilung der Mengen U und D in Untermengen erfolgt für beide Mengen identisch. Aus diesem Grund wird im Folgenden beispielhaft die Menge U herangezogen. Alle Erläuterungen, die zur Separation der Menge U erfolgen, können auf die Menge D übertragen werden.

Die Punktmenge U ist aufsteigend nach den Spalten u jedes Punktes $\mathbf{q}(u, v)$ sortiert. Die Punktmenge U wird in die Untermengen U_i unterteilt, wobei

$$\begin{aligned} \bigcup_{i=1, \dots, m} U_i &= U \\ U_i \cap U_k &= \emptyset \end{aligned} \quad (4.3)$$

⁶Die beiden parallelen Seiten eines Trapezes.

gilt.

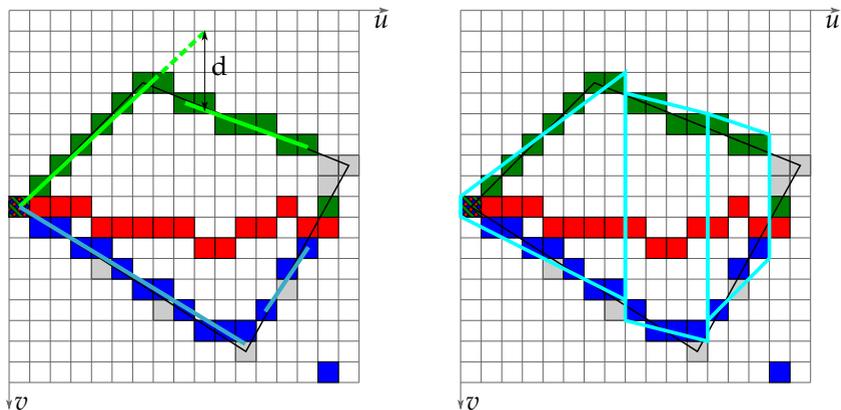
Aufbauend auf der Steigerung der Winkelauflösung des Laserscanners durch pseudo-Messungen (siehe Abschnitt 4.4.2), liegen Lasermessungen für jede Spalte des Bildes vor, in der Teile einer Objektoberfläche aus dem gemeinsamen Sichtbereich abgebildet sind. Spalten ohne korrespondierende Lasermessungen, die von Spalten mit Laserprojektionen flankiert werden, weisen auf zwei getrennte Objektoberflächen hin.

Den Ausgangspunkt jeder Untermenge U_i bilden zwei Pixel aus benachbarten Spalten u und $u + 1$. Diese Pixel bilden die Startlösung einer möglichen Geradengleichung. Liegt in U ein weiteres Element (Aspirant) vor, welches der Spalte $u + 2$ zugeordnet werden kann, wird deren Zugehörigkeit zu der aufgestellten Geradengleichung geprüft. Die Prüfung erfolgt etwa über eine Abstandsbewertung d zwischen der vorliegenden Geraden und dem Aspiranten. Beim Einhalten einer vorab festgelegten Grenze (maximaler Abstand) wird der Aspirant U_i hinzugefügt und die Geradengleichung mittels einer Regression angepasst.

Im Fall, dass der Abstand zwischen dem Aspiranten und der prognostizierten Geraden die festgelegten Grenzen verletzt, wird die Untermenge U_i geschlossen und der Aspirant bildet das erste Element von U_{i+1} .

Kann für einen Punkt $\mathbf{q}(u, v)$, der Teil einer Untermenge U_i ist, kein Nachfolger in der Spalte $u + 1$ in der Menge U gefunden werden, wird U_i geschlossen. Der auf das Element $\mathbf{q}(u, v)$ folgende Aspirant $\mathbf{q}(u + n, w)$ bildet das erste Element von U_{i+1} .

Schlussendlich enthält jede Untermenge U_i und D_j die Datenbasis einer Regressionsgeraden sowie deren Start- und Endpunkt. Eine Visualisierung der bestimmten Gruppierung und der resultierenden Geraden ist in Abbildung 4.11a zu sehen. Die Kantenpixel der Mengen U und D werden jeweils zu zwei Kanten zusammengefasst. Für die Menge U wurde die Abstandsgrenze d eingezeichnet, durch die die Separation erfolgt. Je ein Pixel der Ausgangsmengen U und D zeigt keine Zugehörigkeit zu einer Geraden. Aufgrund des verwendeten Abstandskriteriums wurden diese Pixel keiner bestehenden Untermenge hinzugefügt. In der jeweiligen Untermenge, in welcher sie sich befinden, sind sie jeweils der einzige Punkt, was das Aufstellen einer Geradengleichung nicht ermöglicht. Jede der Untermengen U_i und D_j , die mittels der Separation bestimmt wurden, dient als Schenkel des zu erstellenden trapezförmigen Segmentes, wobei die Grundkanten parallel zur Suchrichtung der Kantenpixel verlaufen. Für jedes Segment wird ein Schenkel durch eine der Mengen U_i und der zweite durch eine der Mengen D_i gestellt. Die Spaltenposition für die Start- und Endpunkte der beiden Mengen müssen identisch sein, um die Bedingung, die durch die Lage der Grundkanten aufgestellt ist, zu erfüllen. Diese geforderte Si-



(a) Kombination einzelner Kantenpixel zu zusammenhängenden Geraden.

(b) Kombination der gefunden Geraden zu trapezförmigen Segmenten

Abbildung 4.11: Erstellung von Segmenten auf Basis von Kantenpixeln.

tuation kann in Abbildung 4.11a nicht beobachtet werden.

Um die angestrebte Situation herbeizuführen, werden die Untermengen U_i und D_j erneut geteilt. Eine Untermenge U_i wird etwa in zwei Abschnitte separiert, wobei die Geradengleichung für beide Abschnitte identisch ist. Allein die Start- und Endpunkte werden so angepasst, dass die Grundkanten wie gefordert parallel zu einer der Koordinatenachsen der Bildebene verlaufen. Die Position der Separation wird durch den Start- bzw. Endpunkt der Menge D_j festgelegt.

Aufbauend auf den bestimmten Untermengen U_i und D_j bzw. den enthaltenden Geraden (siehe Abbildung 4.11a) und der nachfolgenden Teilung der Geraden erfolgt die Bestimmung der trapezförmigen Segmente. Eine beispielhafte Darstellung der Segmente wird durch Abbildung 4.11b skizziert.

4.4.2 Anpassung der LiDAR-Auflösung

Die Segmentierung innerhalb der kooperativen Fusion setzt voraus, dass die Messungen eines Scans, welche auf die Bildebene projiziert werden, auf dieser eine approximierte Schnittgerade \mathbf{g}'_{OA} darstellen. Darüber hinaus wird davon ausgegangen, dass für alle Zeilen bzw. Spalten, welche von \mathbf{g}'_{OA} geschnitten werden, mindestens eine Lasermessung vorliegt. Dies

kann auf Grundlage unterschiedlicher Auflösungen von Kamera und Laserscanner für die Rohdaten nicht sichergestellt werden.

Unterschiedliche Sensorauflösung

Liegt aufgrund von unterschiedlichen Winkelauflösungen zwischen Kamera und Laserscanner nicht für jede Spalte des abgebildeten Objekts eine projizierte Lasermessung vor, ist der Aufbau einer Kontur, die das Segment hinreichend abtrennt, nicht möglich (siehe Abbildung 4.13a). Aus diesem Grund ist eine Anpassung der Auflösungen aneinander notwendig. In dem hier vorgestellten Ansatz wird die Auflösung eines geringer auflösenden LiDAR an die Auflösung einer Kamera vorgenommen.

Aktuell verfügbare Laserscanner weisen oft eine geringere Winkelauflösung als handelsübliche⁷ Kameras auf. Dieser Unterschied führt dazu, dass nicht jede Spalte, die vom approximierten Schnitt g_{OA} durchlaufen wird, eine Lasermessung aufweist. Um eine Separation des Segmentes, welches korrespondierende Lasermessungen aufweist, von weiteren Segmenten auf der Bildebene zu ermöglichen, ist es notwendig, dass auch für diese Spalten die Ober- und Untergrenzen des Segmentes identifiziert werden.

Pseudomessungen

Das Ergebnis einer Laserentfernungsmessung entspricht je nach Art der Echoauswertung, der Entfernung zwischen Laserquelle und der vom Laser beleuchteten Fläche (siehe Abschnitt 2.3.2). Die Größe der beleuchteten Fläche ist durch den Fernfelddivergenzwinkel und die Entfernung zwischen Strahlenquelle und Objektoberfläche bedingt (siehe dazu Abschnitt 2.3.2). Die Größe der durch den Laser beleuchteten Fläche wird in dem bestimmten Ergebnis nicht berücksichtigt. Es wird ausschließlich die Entfernung für einen mathematischen Strahl zum Zentrum der Fläche angegeben.

Die Projektion der ermittelten Position auf die Bildebene ergibt schlussendlich ein korrespondierendes Pixel. Eine Projektion der Fläche des Strahlenquerschnitts, der sich durch die Fernfelddivergenz $w(z)$ ergibt (vergleiche Gleichung 2.15), kann allerdings durchaus größer als ein Pixel ausfallen. Bedeckt die Projektion des Strahlenquerschnitts auf der Bildebene weitere Pixel, die von dem ursprünglichen korrespondierenden Pixel $q_I(u, v)$

⁷Handelsübliche Laserscanner zur Freiraumüberwachung besitzen eine Winkelauflösung von 0.125° bis 0.5° [37, 83, 93]. Eine exemplarische Kamera mit einer 640×480 Pixelauflösung und einer Optik mit 60° Öffnungswinkel ermöglicht eine Winkelauflösung von ca. 0.1°

abweichen, können für diese Pixel ebenfalls korrespondierende Lasermessungen angenommen werden. Die Übernahme der Laserentfernungsmessungen erfolgt ausschließlich für Pixel in benachbarten Spalten bei konstantem v . Visualisiert wird dieser Zusammenhang in Abbildung 4.12.

Der Ansatz „Entfernungsmessungen für benachbarte Pixel zu übernehmen“, deren Rückprojektionen innerhalb des Laserquerschnitt liegen, basiert auf der Annahme, dass der Laser eine „first Echo“-Auswertung verwendet (siehe Abschnitt 2.3.2). Damit ist garantiert, dass kein Bestandteil der Fläche, die von dem Laser beleuchtet wird, eine geringere Entfernung aufweist als das gelieferte Ergebnis. Eine Objektfläche, die einem Pixel mit korrespondierender pseudo-Lasermessung zugeordnet werden kann, weist eine gleiche bzw. höhere Entfernung zum Laser auf, als die Lasermessung auf der sie basiert.

Lasermessungen, die nicht direkt vom Laser generiert werden, sondern

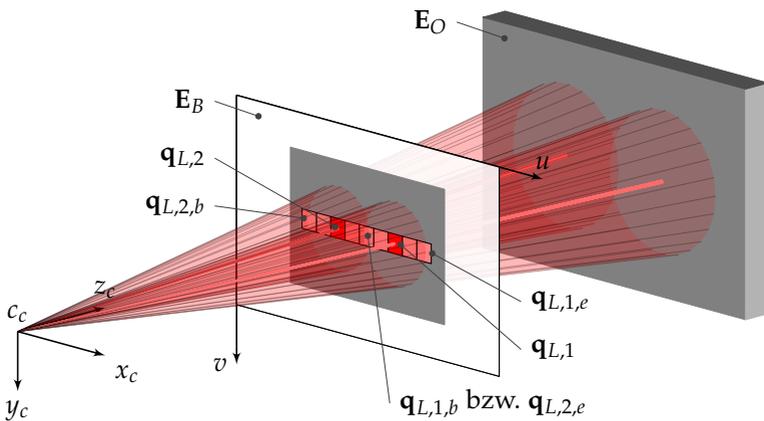
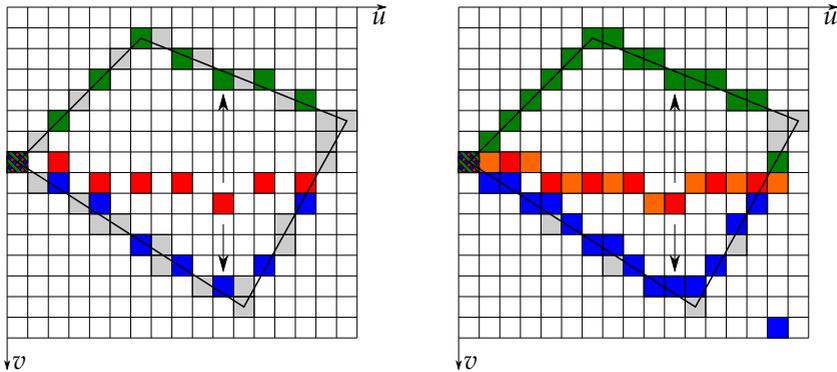


Abbildung 4.12: Berücksichtigung der Laserspotfläche zur Ermittlung von pseudo-Lasermessungen auf der Bildebene.

auf der Ausdehnung des Lasers beruhen, werden im Folgenden als pseudo-Lasermessungen bezeichnet. Jedes Pixel, welches von dem approximierten Schnitt g_{OA} berührt wird, wird bezüglich seiner Position im projizierten Strahlenquerschnitt einer realen Lasermessung mit einer pseudo-Entfernung versehen. Entsteht durch dieses Vorgehen der Fall, dass ein Pixel in dem Einflussbereich von zwei Laserprojektionen liegt, die auf realen Messungen beruhen, wird auf Grundlage der „first Echo“-Auswertung der geringere Abstand zum Sensor verwendet.

Anwendung von Pseudomessungen für die Konturdetektion

Unter der Annahme, dass der Laserscanner die Umwelt derart abtastet, dass zwischen den Laserstrahlen keine Lücken entstehen, $\gamma \geq \Delta\Theta_l^8$, wird jede Spalte, die von einem Schnitt \mathbf{g}_{OA} berührt wird, mit einer auf realen oder pseudo-Messungen beruhenden Projektion versehen. Folglich werden für jede dieser Spalten die Segmentgrenzen ober- und unterhalb von \mathbf{g}_{OA} bestimmt. Das durch die Berücksichtigung der Fernfelddivergenz generierte Resultat in Abbildung 4.13b entspricht damit wieder nahezu dem Resultat aus Abbildung 4.10b, bei der die Winkelauflösung von Kamera und Laser übereinstimmen. Ein Unterschied kann bei der Position der projizierten Lasermessungen ausgemacht werden, die nicht für alle Pseudo-Messungen denen in Abbildung 4.10b entspricht. Dieser Unterschied hat in diesem Beispiel keinen Einfluss auf die Bestimmung der Kontur.



(a) Suche bei ausschließlicher Nutzung der Lasermessungen

(b) Suche mittels realer Lasermessungen und pseudo-Messungen.

Abbildung 4.13: Identifikation von Konturpixeln mittels Lasermessungen, welche die halbe Winkelauflösung der Kamera aufweisen.

Pseudo-Lasermessungen werden nur für den aktuellen Scan \mathbf{A}_t bestimmt und dienen ausschließlich der Suche nach den Segmentgrenzen. In den weiteren Verarbeitungsschritten finden sie keine Berücksichtigung.

⁸In der Praxis ist die Winkelauflösung und die Fernfelddivergenz eines Lasers vom Hersteller oftmals so gewählt, dass keine Lücken zwischen den einzelnen Strahlen entstehen und somit keine Hindernisse übersehen werden [37, 84, 93].

4.4.3 Oberflächenapproximation

Die Segmentierung präsentiert einen Ansatz, der mittels einer Kombination aus Kantendetektion und eines auf die Bildebene projizierten Laserscans A_t , Segmente auf der Bildebene lokalisiert.

Bei der Konzeptionierung der Sichtfelderweiterung (siehe Abschnitt 4.2) wird die Annahme getroffen, dass ein Segment auf der Bildebene die Projektion einer planaren Objektoberfläche im Objektraum ist. Nach Abschluss der Segmentierung liegen die Segmente allein auf der Bildebene vor. Um aus ihnen Approximationen von Objektoberflächen herzuleiten, ist eine Transformation der Segmente in den Objektraum notwendig. Die Bestimmung von Lage und Position für jedes Segment soll mittels korrespondierender Lasermessungen erfolgen. Die Erläuterungen zum Konzept der Sichtfelderweiterung zeigen (vergleiche Abschnitt 4.2), dass die Nutzung einer einzelnen Lasermessung bzw. einer kollinearen Punktmenge eine Transformation der Segmente aus dem Bildraum in den Objektraum nur bedingt ermöglicht.

Die Lasermessungen, welche zur Identifikation der Segmente verwendet wurden, liegen innerhalb einer Scanebene und sind pro Segment kollinear, da E_A die planaren angenommenen Objektoberflächen jeweils in einer Geraden schneiden.

Für eine Bestimmung von Lage und Position eines Segmentes, die frei von Annahmen bzgl. deren Lage gegenüber der Bildebene ist, ist es notwendig, den Segmenten weitere Lasermessungen hinzuzufügen, wobei die Menge der Lasermesspunkte eines Segmentes linear unabhängig sein muss.

Die Bereitstellung linear unabhängiger Lasermessungen erfolgt mittels der Kombination mehrerer Scans des monogonen Scanners beginnend beim Zeitpunkt $t - n$ bis inklusive des aktuellen Scans A_t zu einer Punktwolke $C_{A,t-n:t}$. Ausgehend von einer vorausgesetzten Eigenbewegung des Sensorträgers während der Zeitspanne $t - n$ wird angenommen, dass die Scanebenen E_A der Scans nicht in einer Ebene liegen und die Punkte innerhalb $C_{A,t-n:t}$ linear unabhängig verteilt sind. Für die Kombination der Messungen mehrerer Scans zu einer Punktwolke werden die Messungen von der Bewegung, welche das UA während des Zeitraumes $n - t$ vollführt, in der Vorverarbeitung bereinigt (siehe Abschnitt 4.3.2).

Die Elemente $C_{A,t-n:t}$ werden ebenfalls auf die Bildebene projiziert. Abhängig von der vorausgehenden Bewegung des Sensorträgers verteilen sie sich über die Bildebene. Liegen die Bildkoordinaten einer Lasermessung innerhalb der Grenzen eines Segmentes, werden die entsprechenden Objektkoordinaten dem Segment zu geordnet. Ein Segment besitzt anschließend neben seinen Grenzen im Bildraum eine Punktwolke C_{S_s} von Laser-

messungen im Objektraum.

Regressionsebene

Unter der Bedingung, dass die Anzahl der in C_{S_j} enthaltenden Messungen größer drei ist, und dass mindestens drei dieser Messungen linear unabhängig im Objektraum verteilt sind, wird auf Basis von C_{S_j} eine Ebene approximiert. Die hergeleitete Ebene repräsentiert die angenommene planare Objektoberfläche im Objektraum, auf der das Segment S_j liegt. Die Approximation der Ebene erfolgt mit einem Regressionsansatz unter Verwendung einer Hauptkomponentenanalyse [51].

Mittels der Hauptkomponentenanalyse werden die drei Linearkombinationen bzw. Hauptkomponenten der Punktwolke C_{S_j} gesucht. Zur Herleitung einer Ebene werden die beiden Linearkombinationen mit der größten Varianz zur Beschreibung der beiden Richtungsvektoren verwendet. Die dritte der Linearkombinationen gibt die Varianz der Punktwolke zur Ebene an. Im Folgenden wird speziell auf die Herleitung der Ebene eingegangen. Auf Informationen, die aus der bereitgestellten Varianz gewonnen werden, wird in Abschnitt 4.5.1 näher eingegangen.

Zur Herleitung der Linearkombinationen wird C_{S_j} , welche $n > 3$ Messungen \mathbf{p}_m enthält, bezüglich ihres Erwartungswertes

$$\mathbf{m} = \frac{1}{n} \sum_{m=1}^n \mathbf{p}_m \quad (4.4)$$

normiert. Die normierten Messungen werden zu einem Vektor

$$\mathbf{V} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n] - \mathbf{m} \quad (4.5)$$

zusammengefasst. Die Kovarianzmatrix

$$\text{Cov}(\mathbf{V}) = \mathbf{V}\mathbf{V}^T \quad (4.6)$$

ist symmetrisch und positiv definiert. Mittels der Eigenwerte λ_i von $\text{Cov}(\mathbf{V})$ lassen sich die Eigenvektoren \mathbf{v}_i herleiten. Die Eigenwerte geben Auskunft über die Streckung von C_{S_j} entlang der Hauptkomponenten. Die Eigenvektoren beschreiben die Orientierung der Hauptkomponenten. Die Eigenvektoren \mathbf{v}_1 und \mathbf{v}_2 , welche auf den beiden Eigenwerten λ_1 und λ_2 basieren, wobei $\lambda_1 > \lambda_2 > \lambda_3$ gilt, beschreiben die Orientierung der beiden umfangreicheren Streckungen von C_{S_j} . Eine Ebene kann durch zwei linear unabhängige Vektoren definiert werden. In normierter Form entsprechen \mathbf{v}_1 und \mathbf{v}_2 den zwei linear unabhängigen Richtungsvektoren einer Ebene.

Die Parameterdarstellung der Ebene $E_{O,j}$, die dem Segment S_j zugeordnet ist, wird durch

$$E_{O,j} : \mathbf{m} + s|\mathbf{v}_1| + t|\mathbf{v}_2| \quad (4.7)$$

beschrieben. Die Ebene E_s ist infinit. Ein Einsatz von $E_{O,j}$ als Repräsentation einer Objektfläche setzt eine Begrenzung voraus. Eine Begrenzung wird durch die Rückprojektion der Eckpunkte von S_j von der Bildebene auf $E_{O,j}$ erreicht. Auf Basis der intrinsischen Parameter kann für jeden Eckpunkt des Polygons (Trapez), welches das Segment S_j beschreibt, jeweils ein Strahl bestimmt werden, der unter Annahme einer aberrationsfreien Projektion die Ebene $E_{O,j}$ durchstößt. Die entstehenden Schnittpunkte bilden die approximierten Objektpunkte der bildseitigen Polygonecken. Eine Verknüpfung der Objektpunkte analog zu den korrespondierenden Bildpunkten des Segments ermöglicht die Begrenzung von $E_{O,j}$ auf den Teil, der tatsächlich eine mögliche Objektfläche beschreibt. Der finite Teil von $E_{O,j}$, welcher von der Projektion von S_j eingefasst wird, wird im weiteren als $\hat{E}_{O,j}$ bezeichnet. Die Polygone $\hat{E}_{O,j}$ bilden das Ergebnis der kooperativen Fusion und beschreiben die gesuchten planaren dreidimensionalen Oberflächenapproximationen. Abbildung 4.14 zeigt das Resultat der

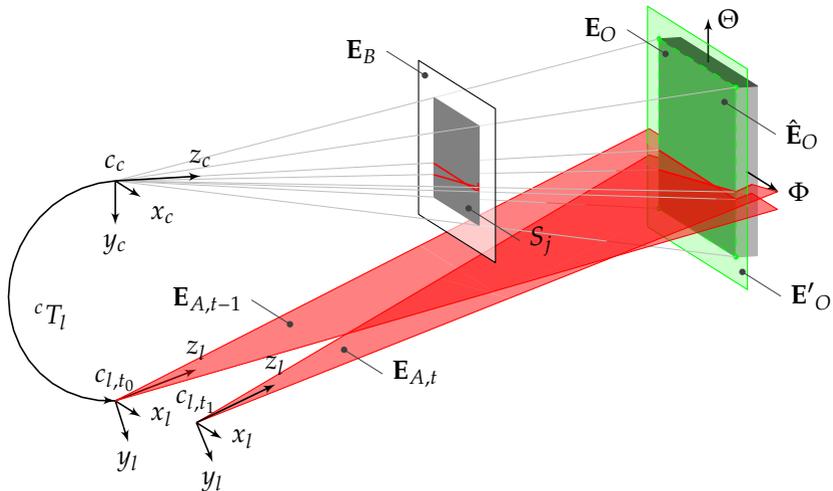


Abbildung 4.14: Planare Oberflächenapproximation als Ergebnis der kooperativen Fusion.

kooperativen Fusion anhand eines planaren Objektes. Die rekonstruierte Oberfläche $E_{O,j}$ entspricht in diesem Fall der tatsächlichen Objektfläche

che $E_{O,j}$.

Kombination von identischen Ebenen

Die Erzeugung der Regressionsebene im Objektraum setzt eine Mindestanzahl von Elementen $\mathbf{p} \in C_{S_j}$ pro Segment S_j voraus. Darüber hinaus sollten die Elemente \mathbf{p} möglichst gleichmäßig über die Fläche des Segmentes auf der Bildebene verteilt sein. Wie viele Messungen innerhalb eines Segmentes platziert sind und wie diese verteilt sind, ist von Größe und Form des Segmentes, der Winkelauflösung des Lasers und der vom UA ausgeführten Bewegung abhängig.

Für eine Approximation der Segmentebene sind mindestens drei Lasermessungen \mathbf{p} notwendig. Die Lage von Segmenten, die eine geringere Anzahl von korrespondierenden Messungen enthalten, können nicht eindeutig bestimmt werden und müssten verworfen werden. Dieses Vorgehen führt allerdings zu einem Verlust von Informationen.

Die in Abschnitt 4.4.1 dargelegte Situation legt nahe, dass die Kanten von Segmenten auch auf Texturen und nicht ausschließlich auf Objektkanten beruhen. Segmente, welche auf Textur zurückzuführen sind, weisen im Allgemeinen kleinere oder gleichgroße Abmessungen auf, als die Objektoberfläche $E_{O,j}$, auf der sich die Textur befindet. Es besteht die Möglichkeit, dass sich auf einer Objektoberfläche mehrere Segmente S_j befinden, deren Trennung auf einer Textur beruht.

Daraus folgt, unter der Annahme einer ebenen Objektoberfläche, dass eine Kombination der korrespondierenden Lasermessungen \mathbf{p} aller Segmente, die auf einer Objektoberfläche liegen, eine einzige Ebene beschreibt [50].

Die Parameter der Ebenen, die aus einer derartigen Kombination entstünden, könnten auch für wenige aussagekräftige Segmente verwendet werden. Ein Verwerfen dieser Segmente wäre damit nicht mehr notwendig. Voraussetzung für ein derartiges Vorgehen ist die Bestimmung der Segmente, welche auf einer Objektoberfläche $E_{O,j}$ liegen. Diese soll ausgehend von $E_{O,j}$ erfolgen. Bei der Bestimmung einer $E_{O,j}$ aus mehr als drei linear unabhängigen Messungen \mathbf{p} weisen diese eine Varianz zu $E_{O,j}$ auf. Fügt man einer Menge an Messungen, welche ein signifikantes Segment beschreiben, weitere Messungen aus einem nicht aussagekräftigen Segment hinzu, sollte eine Veränderung der Varianz zu beobachten sein. Das Ausmaß dieser Änderung wird als Indikator verwendet, ob sich Messungen unterschiedlicher Segmente auf ein und derselben $E_{O,j}$ befinden. Liegen die Änderungen der Varianz unter einem vorab definierten Schwellwert, etwa der Varianz der Laserrohdaten, wird davon ausgegangen, dass beide Segmente auf der identischen $E_{O,j}$ beruhen. In dem Fall, dass die Segmente

auf eine Ebene zurückzuführen sind, werden die Punktmengen vereinigt und die resultierende Regressionsebene bildet für beide Segmente die Prädiktion im Objektraum.

Der Vergleich erfolgt zwischen allen Segmenten. Bleiben nach der Durchführung der Vergleiche uneindeutige Segmente bestehen, werden diese schlussendlich doch entfernt.

Dieses Vorgehen führt besonders bei kleineren Segmenten zu einer Steigerung der Aussagekraft bzw. ermöglicht überhaupt erst die Bereitstellung einer Ebene [50].

4.5 Konkurrierende Fusion

Dreidimensionale Messungen, die direkt von Sensoren oder Fusionsergebnissen, etwa Stereoskopie bzw. dem Vorgehen der vorliegenden Arbeit, bereitgestellt werden,

- zeigen nur einen Teilbereich der umgebenden Umwelt,
- sind mit Fehlern behaftet, etwa durch Sensorrauschen,
- enthalten dynamische Objekte, die nicht dauerhaft kartographiert werden sollen,
- oder stellen nur Prognosen dar [89].

Im Hinblick auf die präsentierte Fusion, beschränken die Sichtfelder der Ausgangssensoren, Kamera und Laserscanner, die Sicht auf die Umwelt. Die Ergebnisse der Fusion können maximal das Sichtfeld der Kamera einnehmen, welches einer spitzen Rechteckpyramide entspricht. Was sich außerhalb dieser Pyramide befindet, kann auf den „ersten Blick“, auf Basis einer einzelnen Momentaufnahme, nicht erfasst werden. Eine Momentaufnahme ist ein Ergebnis der Fusion, welches auf dem Kamerabild eines Zeitpunkts t und mehreren Laserscans basiert. Für Objekte bzw. Hindernisse, deren Größe nicht mit einer Momentaufnahme erfasst werden kann, gilt, dass erst eine Kombination mehrerer iterativer Aufnahmen des Objektes, welche aus unterschiedlichen Posen gemacht wurden, eine Abschätzung des gesamten, belegten Raums ermöglicht.

Beim Aufzeichnen der Sensorrohdaten durch die Sensoren können zahlreiche Quellen ursächlich für falsche oder fehlende Ergebnisse sein. Dazu zählen Sensorrauschen, Messauflösung oder die Gestalt, Form und Material der betrachteten Objekte, welche Einfluss auf das jeweilige Messverfahren haben.

Dynamische Objekte sollen nur bedingt kartographiert bzw. als Hindernisse betrachtet werden. Solange sich die dynamischen Hindernisse im direkten Aktionsradius des unbemannten Vehikels befinden, sollten sie als potenzielle Gefahren klassifiziert werden und der durch sie belegte Raum ebenfalls als dieser gekennzeichnet werden. Findet darüber hinaus eine Bewegung durch das dynamische Objekt statt, ist eine Detektion des freiwerdenden und des neu zu belegenden Raums notwendig. Beides ist mittels Momentaufnahmen nicht möglich, da diese jeweils eine statische Situation zu einem spezifischen Zeitpunkt t zeigen, zu welchem dynamische Objekte ebenfalls statisch erscheinen.

Die Auslegung der bis hier vorgestellten Fusion von monokularen Kamera- und Laserscannerdaten geht von zahlreichen Annahmen aus, die nicht oder nur bedingt durch reale Messungen gestützt werden. Es handelt sich um Approximationen und Prognosen, deren Eintrittswahrscheinlichkeit als gering eingeschätzt werden sollte und mittels einer Momentaufnahme nicht verifiziert werden kann.

Zur Berücksichtigung der dargelegten Schwächen wird in diesem Kapitel eine konkurrierende Fusion mittels einer dreidimensionalen Umweltrepräsentation vorgestellt, welche es erlaubt, einerseits die Unzulänglichkeiten der Sensoren, wie begrenztes Sichtfeld und Messfehler zu reduzieren, und andererseits mit einer dynamischen Umwelt umzugehen oder Prognosen bzw. Approximationen über die Zeit zu stärken oder zu verwerfen. Zur Darlegung dieses Vorgehens wird auf die Resultate der kooperativen Fusion aus Abschnitt 4.4 eingegangen. Anschließend folgt eine Erweiterung des inversen Sensormodells (siehe Abschnitt 2.3.2), um die Approximationen der kooperativen Fusion mit einer Signifikanz zu versehen, mit dem Ziel, die probabilistische Inklusion in der dreidimensionalen Umweltrepräsentation zu ermöglichen. Die Aufarbeitung der Approximationen erlaubt schlussendlich das Eingliedern der kooperativen Fusionsdaten in eine dreidimensionale Umweltrepräsentation und damit das Herausbilden der wahren Objektstruktur.

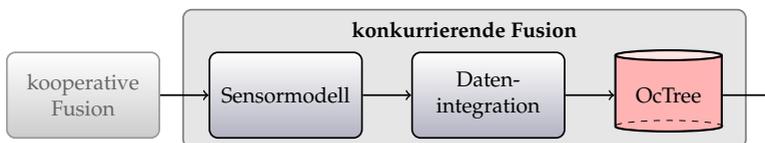


Abbildung 4.15: Blockdiagramm der konkurrierenden Fusion zur Kombination der vorausgehenden Fusionsergebnisse.

4.5.1 Interpretation von Sensordaten

Der Abschnitt 4.4.3 schließt mit der Repräsentation der identifizierten Objektflächen durch Ebenen. Abbildung 4.14 zeigt eine derartige Momentaufnahme, in der zwei Scanebenen $E_{A,t-1}$ und $E_{A,t}$ des Lasers mit einem einzigen Kamerabild fusioniert werden, um eine Ebene E'_O und schlussendlich die Oberfläche \hat{E}_O (grün gestrichelt) zu bestimmen, welche das betrachtete Objekt abbildet. Dieses Vorgehen beschreibt die Objektfläche allerdings nur dann hinreichend korrekt, wenn das Objekt ebenfalls eine planare Oberfläche aufweist. Dies ist keine Voraussetzung bei natürlichen Objekten. Im Resultat entstehen Abweichungen zwischen der prognostizierten Fläche und der wahren Objektfläche, wie Abbildung 4.16 verdeutlicht. Die entstehenden Abweichungen können mit einer ein-

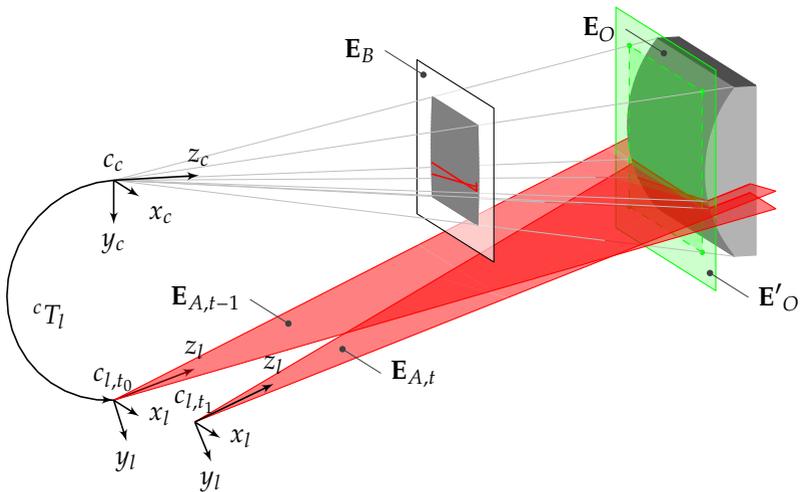


Abbildung 4.16: Abweichung planarer Oberflächenapproximation bei nicht planaren Oberflächen.

zigen Momentaufnahme nicht identifiziert werden. Es ist eine Kombination mehrerer Momentaufnahmen notwendig. Eine direkte Kombination der jeweiligen Flächen als Ganzes wird allerdings nicht nahegelegt, da die Abweichungen über die Flächen unterschiedlich verteilt sein können. Am Beispiel in Abbildung 4.16 ist dies gut ersichtlich. Im Bereich, in dem die Lasermessungen das Objekt abtasten, liegt eine relativ gute Übereinstimmung vor. Je größer der Abstand zu diesem Bereich wird, desto stärker werden die Abweichungen zwischen Objektfläche und Approximati-

on.

Aufbauend auf diesen Beobachtungen wird keine Kombination der jeweils gesamten Flächen angestrebt. Vielmehr sollen diskrete Prädiktionen, die sich innerhalb der bestimmten Ebenen $\hat{\mathbf{E}}_O$ befinden, über mehrere unterschiedliche Momentaufnahmen hinweg gewichtet kombiniert werden. Die Gewichtungsfaktoren werden durch die Lage über der Fläche, der Position einer jeweiligen Prädiktion auf der Fläche und die Eigenschaften der Ausgangssensoren, Kamera und Laserscanner, bestimmt.

Die Diskretisierung von $\hat{\mathbf{E}}_O$ erfolgt auf Basis eines Kugelkoordinatensystems dessen Ursprung c_C das Kamera-System bildet. Dieses Vorgehen wurde gewählt, um das Abbildungsvermögen der Kamera nachzubilden. Jeder bei der Diskretisierung entstehende approximierter Messpunkt $\tilde{\mathbf{p}}$ repräsentiert den Schnittpunkt eines Strahls, der vom Kameraursprung ausgeht, mit einer der Flächen E_O . Die Schrittweite der Diskretisierung wird mit der Winkelauflösung $\Delta\Theta_s$ bzw. $\Delta\Phi_s$ umgesetzt. Die maximal erreichbare Winkelauflösung sowie das Sichtfeld $\Theta_s \times \Phi_s$ entsprechen den Parametern der Kamera.

Das vorgestellte inverse Lasermodell in Abschnitt 2.3.2 ist für die direkte Anwendung auf die Rohdaten p eines die Entfernung messenden Sensors vorgesehen. Die resultierenden Approximationen $\tilde{\mathbf{p}}$ der Diskretisierung basieren auf Regressionen, deren Basis die Rohdaten eines Lasers sind. Ein direkter Zusammenhang zwischen den Rohdaten des Lasers und den Approximationen ist aber nicht gegeben. Um eine Anwendung des inversen Sensormodells ebenfalls für die $\tilde{\mathbf{p}}$ der Sichtfelderweiterung zu ermöglichen wird das grundlegenden Modell aus Gleichung 2.25 um die drei folgenden Aspekte erweitert:

- Abweichungen einer Regressionsebene gegenüber Laserrohdaten
- Entfernung einer Approximation $\tilde{\mathbf{p}}$ zur Lasersicht
- Auflösungsvermögen der Kamera

Abweichungen einer Regressionsebene gegenüber Laserrohdaten

Inwieweit sich die Rohdaten einer mittels Laser vermessenen Oberfläche mit einer approximierten Ebene decken, kann über die Varianz der Ausgangsmessungen gegenüber der Ebene ermittelt werden. Die Varianz beschreibt das Rauschen der Ausgangsdaten orthogonal zur erwarteten Ebene.

Teilmengen der Laserrohdaten, die innerhalb eines Segments auf der Kamerabildebene liegen, stellen die Datenbasen für die Ermittlung der jeweiligen Regressionsebene \mathbf{E}'_O , welche mittels einer Hauptkomponentenana-

lyse bestimmt wird.

Ausgehend von den Erläuterungen in Abschnitt 4.4.3, der Erstellung der Ebene auf Basis der beiden ausgeprägtesten Hauptkomponenten \mathbf{v}_1 und \mathbf{v}_2 , wird die dritte Komponente \mathbf{v}_3 , die orthogonal zu \mathbf{v}_1 und \mathbf{v}_2 orientiert ist, als Ausgangspunkt für die Varianzermittlung der Ebene genutzt. Nach Jolliffe [51, S. 5] lässt sich die Varianz $Var(\lambda_i)$ einer der Hauptkomponenten wie folgt bestimmen:

$$Var(\lambda_i) = \lambda_i^T \cdot Cov(\mathbf{V}) \cdot \lambda_i. \quad (4.8)$$

Die Varianz der dritten Hauptkomponente $Var(\lambda_3)$ gibt folglich Auskunft darüber wie stark die Laserrohdaten orthogonal zur Ebene ausgelenkt sind. Je nach Umfang der Varianz $Var(\mathbf{E}'_O) = Var(\lambda_3)$ wird die Abbildung der Objektoberfläche durch \mathbf{E}'_O in Frage gestellt. Anders ausgedrückt reduziert sich mit zunehmender $Var(\mathbf{E}'_O)$ die Wahrscheinlichkeit, dass eine berechnete Approximation $\tilde{\mathbf{p}}$, die auf \mathbf{E}'_O basiert, ein Pendant an gleicher Position hat, das einer realen, tatsächlichen Sensormessung entspricht. Eine Reduktion der Wahrscheinlichkeit führt wie die Auswirkungen des physikalischen Sensormodells zu einer Reduktion der Signifikanz.

Ausgehend von der Standardabweichung σ_l einer Lasermessung gegenüber einer bekannten Oberfläche wird die Varianz $Var(\mathbf{E}'_O)$ der Regressionsebene in dem inversen Sensormodell berücksichtigt (siehe Abschnitt 2.3.2). Für den idealen Fall, dass eine ebene Objektoberfläche mittels Lasermessungen vermessen wurde, ergibt sich eine Regressionsebene, die die Objektoberfläche genau nachbildet. Die Standardabweichung von \mathbf{E}'_O ist $\sigma_E = \sqrt{Var(\mathbf{E}'_O)} \leq \sigma_l$. Jeder Abweichung der Objektoberfläche von einer Ebene führt zu einem Anstieg von $Var(\mathbf{E}'_O)$ bzw. σ_E . Ausgehend von der minimalen angenommenen Standardabweichung ergibt sich für σ_E folgende Bedingung:

$$\sigma_E = \begin{cases} \sigma_l & \text{für } \sigma_E \leq \sigma_l, \\ \sigma_E & \text{sonst.} \end{cases} \quad (4.9)$$

Unter Verwendung von σ_E in Abschnitt 2.25 ergibt sich das inverse Sensormodell für einen approximierten Punkt $\tilde{\mathbf{p}}$ auf der Regressionsebene wie folgt:

$$p_{obj}(Z, Z_{\tilde{\mathbf{p}}}) = \left(\eta_l \frac{1}{\sigma_E \sqrt{2\pi}} + p_{unknown} - p_{base}(Z, Z_{\tilde{\mathbf{p}}}) \right) e^{\left(-\frac{(Z - Z_{\tilde{\mathbf{p}}})^2}{2\sigma_E^2} \right)}. \quad (4.10)$$

Visualisiert wird das Modell für verschiedene Varianzen $Var(\mathbf{E}'_O)$ und der daraus resultierenden Signifikanz in Abbildung 4.17.

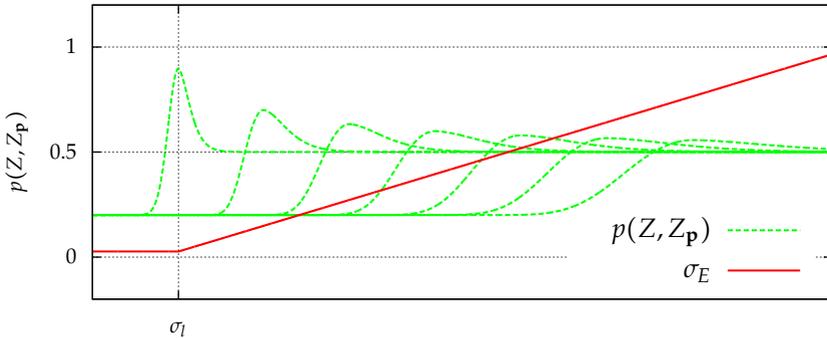


Abbildung 4.17: Signifikanzen bei unterschiedlichen Varianzen von approximierten Objektoberflächen gegenüber deren Rohdaten.

Entfernung einer Approximation $\tilde{\mathbf{p}}$ zur Lasersicht

Auf Basis der Varianz der Laserrohdaten gegenüber der Regressionsebene E'_O lassen sich Differenzen zwischen der prognostizierten Ebene und der Objektoberfläche, innerhalb der durch den Laserscanner generierten Punktwolke aufzeigen. Die Ausdehnung der Punktwolke ist vom Sichtfeld des Lasers und dessen Bewegung gegenüber dem lokalen Weltkoordinatensystem, in dem sich die Objekte befinden, bedingt. Das Sichtfeld des Laserscanners ist orthogonal zur Scanebene gegenüber dem der Kamera eher klein. Diese Situation führt zu der Annahme, dass für ein, auf der Bildebene, identifiziertes Segment, nur für einen kleinen Bereich der Segmentfläche korrespondierenden Lasermessungen vorliegen.

Bei einer alleinigen Verwendung der in Abschnitt 4.5.1 vorgestellten Reduktion der Signifikanz auf Grundlage der Varianz zwischen Laserrohdaten und Regressionsebene, ergibt sich für ausgewählte Szenarios keine Anpassung der Signifikanz, obwohl die Regression die Objektoberfläche ausschließlich in einem kleinen Bereich korrekt widerspiegelt. Abbildung 4.18 präsentiert ein Beispielszenario, in dem ein Objekt mit einer zylindrischen Form detektiert wird. Die erstellte Regressionsebene E'_O weicht von der Objektoberfläche E_O außerhalb der Lasersicht ab. Approximationen $\tilde{\mathbf{p}}$, die auf E'_O beruhen, weichen von einer Messung \mathbf{p} , die auf der tatsächlichen Objektoberfläche liegen, ab. Das angestrebte Resultat ist daher eine Reduktion der Signifikanz von $\tilde{\mathbf{p}}$. Der Bereich, für den Lasermessungen vorliegen, verläuft quasi parallel zur Rotationsachse des Zylinders. Der Radius des Zylinders ist entsprechend groß und die Varianz der Lasermessung

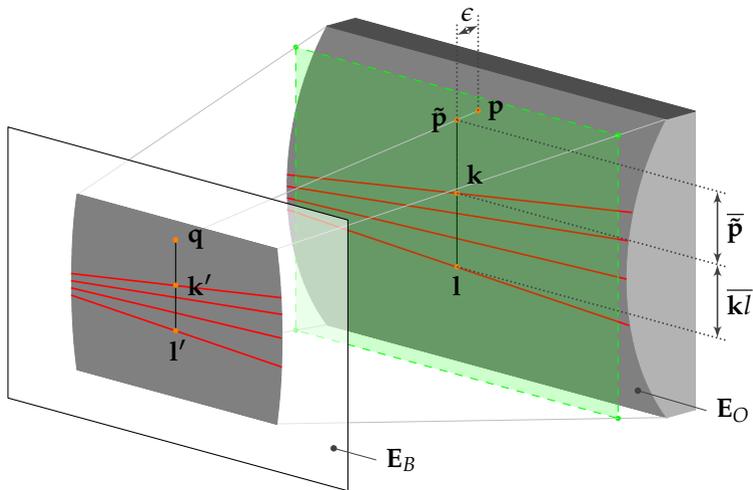


Abbildung 4.18: Reduktion der Signifikanz auf Grundlage der Entfernung einer Approximation \tilde{p} zur Lasersicht.

ist $\text{Var}(\mathbf{E}'_O) \leq \sigma_l$. Durch die runde Form des Objektes sind keine formgebenden Kanten gegeben, welche das Segment begrenzen könnten. Allein durch unterschiedliche Lichtintensitäten auf der Objektoberfläche, welche durch die unterschiedlichen Orientierungen der Oberflächennormalen zur Lichtquelle entstehen, können sich Kanten bei der Segmentierung bilden. Ober- und unterhalb des Bereiches, der von den Lasermessungen abgedeckt wird, weicht die approximierte Ebene von der Objektoberfläche ab. In welchem Umfang die Abweichungen vorliegen, kann aus den Kameradaten nicht ermittelt werden.

Es wird die These aufgestellt, dass die Wahrscheinlichkeit einer Übereinstimmung zwischen einer Approximation, die auf einer Regression basiert, und einer Messung, die auf der Objektoberfläche liegt, sinkt, je weiter sich die Approximation von dem Bereich entfernt, für den korrespondierende Lasermessungen vorliegen.

Für die Berücksichtigung der These im inversen Sensormodell der Fusion soll die Ausdehnung der Lasermessungen auf einer Segmentebene in Bezug gestellt werden zu dem Abstand, der sich zwischen einer betrachteten Approximation und den Lasermessungen aufspannt. Das Vorgehen ist dafür wie folgt: Die auf die Bildebene projizierten Lasermessungen werden mit einem Trapez umgeben, dessen Grundkanten parallel zu den Grundkanten des Segments verlaufen und dessen Fläche minimal ist. Eine Pro-

jektion der Approximation \mathbf{q} auf der Bildebene ist Ausgangspunkt einer Geraden, die parallel zu den Grundkanten des Segments verläuft und beide Schenkel des Trapezes schneidet, das die projizierten Lasermessungen umrahmt. Die dabei entstehenden Schnittpunkte \mathbf{k}' und \mathbf{l}' werden auf die Regressionsebene projiziert und bilden mit der Approximation die Grundlage für den Vergleich. Visualisiert sind die Punkte in Abbildung 4.18.

Der Vektor $k\tilde{p}$ verläuft von Punkt \mathbf{k} des Trapezschenkels zur Approximation $\tilde{\mathbf{p}}$. Ein zweiter Vektor kl zeigt vom Punkt \mathbf{k} zum Punkt \mathbf{l} . Der Abstand $|\overline{kl}|$ wird zu $|k\tilde{p}|$ ins Verhältnis gesetzt. Das Verhältnis gibt Auskunft über die Entfernung von $\tilde{\mathbf{p}}$ bezüglich der Ausdehnung der auf E_s befindlichen Lasermessungen entlang der betrachteten Geraden \tilde{pl} .

Liegt die Approximation innerhalb des Trapezes, das die Lasermessungen umrahmt, wird die Annahme getroffen, dass Abweichungen zwischen der Regressionsebene und der Objektoberfläche durch die Varianz der Regressionsebene (Abschnitt 4.5.1) berücksichtigt sind. Es ist dann keine weitere Betrachtung notwendig. Befindet sich die Approximation außerhalb des umrahmenden Trapezes, findet mit zunehmender Entfernung eine Reduktion der Signifikanz statt.

Die Reduktion der Signifikanz wird mittels einer Substitution von η_l durch

$$\eta_d = \begin{cases} \eta_l & \text{für } k\tilde{p} + kl < kl, \\ \frac{|\overline{kl}|}{|k\tilde{p}|} \cdot \eta_l & \text{für } k\tilde{p} + kl > kl \end{cases} \quad (4.11)$$

in Gleichung 2.25 erreicht.

Auflösungsvermögen der Kamera

Neben Abweichungen der Regressionsebene gegenüber der Umwelt, welche auf Lasermessungen zurückzuführen sind, können auch die Kameradaten der Ursprung von Fehlern während der Fusion sein. Eine Berücksichtigung dieser Fehler ist bei der Bewertung der approximierten Messungen deshalb ebenso notwendig. Das Kameramaterial wird zur Identifikation der möglichen Segmentkanten auf der Bildebene herangezogen. Die Durchführung einer Kantendetektion setzt voraus, dass das Bildmaterial grundsätzlich für Bildverarbeitung geeignet ist, d.h. dass Unter- bzw. Überbelichtung ausgeschlossen bzw. detektiert wird und Belichtungsverhältnisse vorliegen, die eine Detektion von Kanten ermöglichen. Eine Bewertung, ob das Bildmaterial unabhängig vom Inhalt geeignet ist, kann u.a. mit der Auswertung des Histogramms erreicht werden. Eine Auswertung des Histogramms findet im Vorfeld der Segmentierung in der Bildvorverarbeitung zur Steigerung der Dynamik statt. Diese Maßnahme hat

das Ziel, leichte Fehlbelichtungen auszugleichen, um Bildmaterial mit einer vertretbaren Qualität bereitzustellen. Umfangreiche Fehlbelichtungen bis hin zum vollständigen weißen bzw. schwarzen Bild können mit diesem Verfahren allerdings nicht aufgewertet werden. Das Eintreten dieser Fälle muss detektiert und kommuniziert werden.

Ein Auftreten dieser extremen Zustände wird für den aktuellen Ansatz ausgeschlossen. Eine Berücksichtigung von milder schweren Über- und Unterbelichtungen wird durch die Vorverarbeitung abgefangen. Aus diesem Grund findet keine Berücksichtigung der Belichtung der Kamera bei der Signifikanzbestimmung einer approximierten Messung statt.

Ein anderer Punkt, welcher bei der Bestimmung der Signifikanz berücksichtigt wird, ist das Auflösungsvermögen der Kamera. Eine Digitalkamera ist mit einem Sensor versehen, der ein Array von Pixeln enthält, das die durchschnittliche Intensität des auftreffenden Lichts als verarbeitbare Maßzahl bereitstellt. Das auftreffende Licht wird von der abgebildeten Umwelt reflektiert und trifft über die Optik auf die jeweiligen Pixel. Mit zunehmender Entfernung zwischen Kamera und Szene reduziert sich das Auflösungsvermögen, was die Unterscheidbarkeit zwischen feinen Strukturen verhindert. Dies kann dazu führen, dass Kanten übersehen und Segmente bestimmt werden, die größer sind als die zugrundeliegenden Objektflächen.

Eine Orientierung wie hoch das Auflösungsvermögen einer Kamera ist gibt die Menge der Pixel, die einer korrespondierenden Fläche im Objekttraum zugeordnet werden können. Allerdings handelt es sich dabei ausschließlich um einen theoretischen Wert. Die verwendete Optik und der Einfluss des Nyquist-Shannon-Abtasttheorem⁹ (z.B. Moiré-Effekt) können zu Verfälschungen führen, die nicht rückgängig gemacht werden können. Beide Einflussfaktoren führen zu einer Reduktion des Auflösungsvermögens.

Ausgehend von der aberrationsfeien Abbildung des Lochkameramodells und der Einhaltung des Nyquist-Shannon-Abtasttheorems lässt sich mittels des Strahlensatzes

$$\frac{a}{Z_a} = \frac{b}{Z_b} \quad (4.12)$$

die Anzahl der Pixel bestimmen, die zur Verfügung stehen, um feine Strukturen auf einer korrespondierenden Fläche zu identifizieren. Für eine quadratische Fläche mit der Kantenlänge $b = 1m$ ergibt sich für eine Entfernung $Z_b = Z$ eine Auflösung von $n = \frac{a[m]}{a_p[m]}$ bei einer Brennweite $Z_a = f[m] =$

⁹Die Frequenz, mit der ein Bild abgetastet wird, muss mindestens doppelt so groß sein wie die maximal im Bild auftretende Frequenz. Ein Unterlassen führt zu unerwünschten Bildverfälschungen.

$f[p] \cdot a_p[m]$. Die Kantenlänge eines Pixels ist a_p .

$$\frac{a_p[m] \cdot n}{f[p] \cdot a_p[m]} = \frac{1[m]}{Z[m]} \Rightarrow n(Z) = \frac{f[p]}{Z[m]} \tag{4.13}$$

Die bereitgestellte Auflösung der Kamera nimmt mit zunehmender Entfernung ab. Im Nahbereich geht die Auflösung weit über das Maß hinaus, was für die Hindernisdetektion benötigt wird. Aus diesem Grund wird eine Grenze Z_{min} bestimmt, oberhalb welcher keine Variation der Signifikanz einer Messung erfolgt.

$$\eta(Z) = \begin{cases} \eta_l & \text{für } Z \leq Z_{min}, \\ \eta_l \cdot \frac{n(Z)}{n(Z_{min})} & \text{sonst.} \end{cases} \tag{4.14}$$

Die Berücksichtigung des entfernungsabhängigen $\eta(Z)$ in Gleichung 2.25 an Stelle von η_l , führt zu einer Variation der Signifikanz für die approximierten Messungen innerhalb des Bereichs des Sichtfeldes für das keine Lasermessungen vorliegen. Visualisiert wird der Rückgang der Signifikanz auf Grundlage von $\eta(Z)$ in Abbildung 4.19.

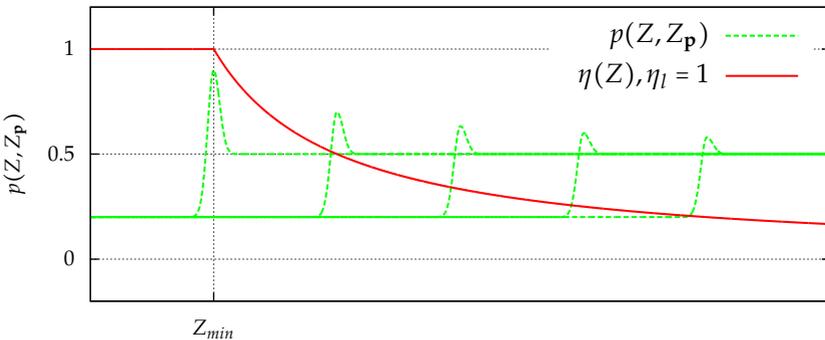


Abbildung 4.19: Signifikanz einer Approximation auf Grundlage des entfernungsabhängigen Kameraauflösungsvermögens.

4.5.2 Rekonstruktion einer Objektoberfläche mittels unabhängiger Prognosen

Aufbauend auf der einmaligen kooperativen Fusion eines Kamerabildes E_B mit mehreren Laserscans $E_{A,t-1}$ und $E_{A,t}$ in Abschnitt 4.4, der Diskretisierung der Objektoberfläche und dem schlussendlichen Einfügen dieser

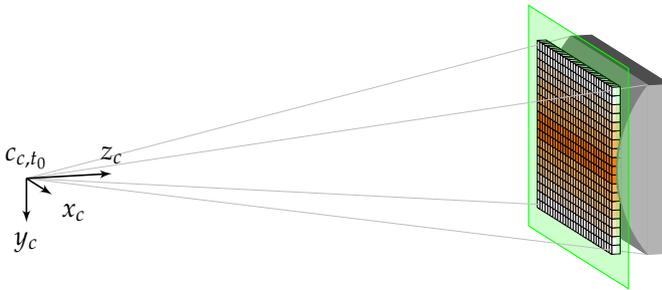


Abbildung 4.21: Diskrete Messungen $\tilde{\mathbf{p}}$ mit Sensormodell in OcTree integriert.

die Abbildung 4.21. Für eine übersichtlichere Visualisierung wird im Folgenden auf die Kameraebene und das Laserkoordinatensystem c_L verzichtet und nur noch das Koordinatensystem c_C dargestellt. Das Kamerakoordinatensystem c_C wird, wie erläutert, als Sensorkoordinatensystem der Sichtfelderweiterung verwendet. Abbildung 4.21 zeigt ein Beispiel, in dem alle Prognosen $\tilde{\mathbf{p}}$ etwa die gleiche Entfernung zu c_C aufweisen und die Varianz der ursprünglichen Lasermessungen gegenüber der resultierenden Regressionsebene kleiner als die Auflösung des OcTrees ist. Das führt dazu, dass der Einfluss der Varianz und der Kameraauflösung auf alle dargestellten Zellen identisch ist und ausschließlich der Abstand zwischen $\tilde{\mathbf{p}}$ und dem Bereich, für den Lasermessungen vorliegen, ausschlaggebend für eine Reduktion der jeweiligen $p(Z, Z_{\tilde{\mathbf{p}}})$ ist.

Ebenso wie in Abbildung 4.20 repräsentiert der OcTree in Abbildung 4.21 ausschließlich die Objektoberfläche, die auf der kooperativen Fusion basiert. Die wahre Gestalt des Objektes wird nicht abgebildet. Es ist allerdings zu erkennen, dass die Zellen des OcTrees, für die keine Lasermessungen vorliegen, mit einer reduzierten Belegungswahrscheinlichkeit versehen sind. Zugleich wird sichtbar, dass die Zellen mit einer reduzierten Belegtheit ebenfalls größere Abweichungen zur erwarteten Objektoberfläche aufweisen.

Wie einleitend beschrieben, genügt die Anwendung einer einzigen Aufnahme der kooperativen Fusion nicht, um die erwartete Objektoberfläche zu erhalten. Vielmehr ist die Kombination mehrerer Aufnahmen innerhalb des OcTrees notwendig, um wie ein Bildhauer die Gestalt des Objekts in mehreren Iterationen herauszuarbeiten. Über die Anwendung der inkrementellen Aktualisierung werden gleiche Belegungen gestärkt und gegensätzliche reduziert (siehe Abschnitt 2.5.2). Die innerhalb des Sensormodells für die kooperative Fusion umgesetzte Reduktion von $p(Z, Z_{\tilde{\mathbf{p}}})$ für die Zel-

len ohne Lasermessungen ermöglicht damit einen schnelleren Wechsel von „belegt“ zu „frei“, was ein gezieltes Herausarbeiten der Objektkontur erlaubt.

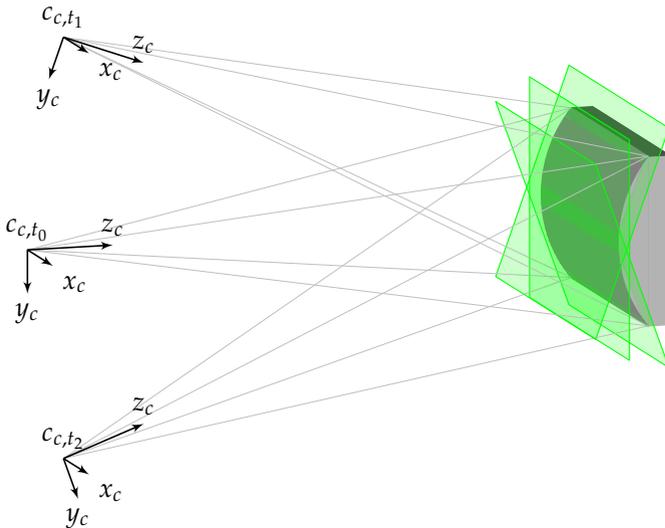


Abbildung 4.22: Betrachtung eines Objektes aus mehreren Blickwinkeln.

Ausgangspunkt für dieses Vorgehen sind mehrere Aufnahmen der erwarteten Objektoberfläche aus unterschiedlichen Posen und, wenn nur eine Sensorkombination zur Verfügung steht, ebenfalls zu unterschiedlichen Zeitpunkten t . Für das hier betrachtete Beispiel werden drei Posen verwendet, die zu den Zeitpunkten t_0 , t_1 und t_2 eingenommen werden. Aus jeder Pose wird eine Prognose generiert, welche die erwartete Objektoberfläche als Ebene prognostiziert. Visualisiert wird diese Situation in Abbildung 4.22. Werden die ebenen Objektoberflächenprognosen aus Abbildung 4.22 jeweils diskretisiert und in den OcTree übertragen, ergibt sich die Situation in Abbildung 4.23. Jede der diskretisierten Approximationen $\tilde{\mathbf{p}}$ repräsentiert einen Strahl, deren Ende ein Hit auf den prognostizierten Ebenen hinterlässt. Bereits belegte Zellen werden geschnitten und aufgrund des iterativen Updates (siehe Abschnitt 2.5.2) in ihre Belegungswahrscheinlichkeit p reduziert. Geringere p werden in Abbildung 4.23 durch hellere Rottöne bei der Färbung der OcTree-Zellen verdeutlicht. Eine Reduzierung von p ist bei den Zellen zu erkennen, welche sich zwischen der erwarteten Objektoberfläche und den jeweiligen Koordinatenursprüngen befinden. Die

Reduktion ist durch das Queren mehrerer Strahlen und ein bereits anfänglich geringes p zu erklären, da die Zellen nicht im Bereich der ursprünglichen Lasermessungen liegen.

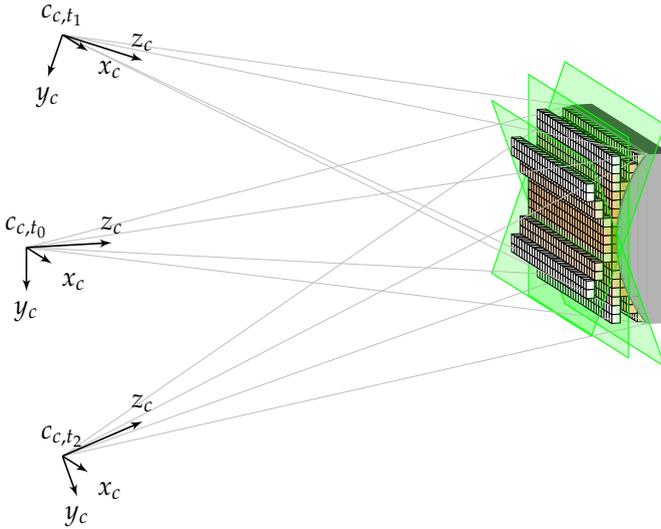


Abbildung 4.23: Integration der diskreten Approximation $\tilde{\mathbf{p}}$ von drei Prognosen in den OcTree.

Mittels mehrerer Iterationen der kooperativen Fusion und der Integration der Daten durch das Bayers Update (siehe Gleichung 2.60) in den OcTree ergibt sich schlussendlich, eine Repräsentation der Objektflächen, die der erwarteten Objektfläche sehr nahe kommt. Siehe dazu Abbildung 4.24.

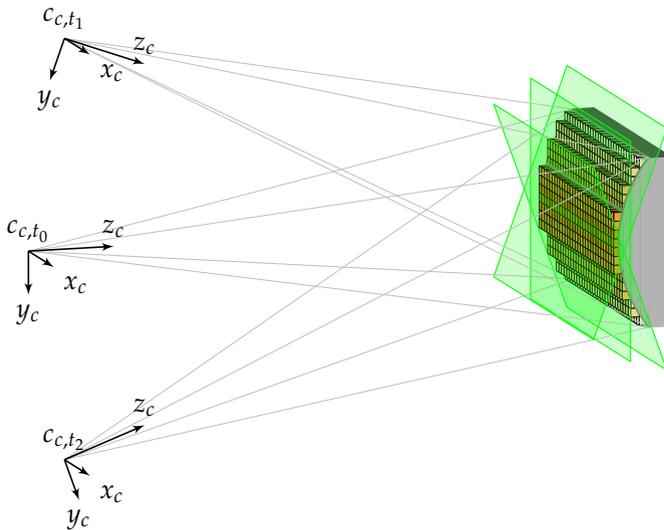


Abbildung 4.24: Rekonstruktion der erwarteten Objektoberfläche durch Integration von n Oberflächenprognosen in den OcTree.

4.6 Validierung der Sichtfelderweiterung

Im Folgenden wird auf den verwendeten Flugversuchsträger und die Sensorhardware sowie die Validierung der vorgestellten Algorithmen eingegangen.

4.6.1 Versuchshardware

Versuchsträger

Für die durchgeführten Versuche wurde ein unbemannter Hubschrauber eingesetzt, dessen Flugleistungen und Nutzlastkapazitäten den Transport von Sensor- und Rechnerhardware ermöglichen, die zur Durchführung der gewünschten Applikationen benötigt wurden. Die Abbildung 4.25 zeigt diesen Hubschrauber, der als „midiARTIS“ bezeichnet wird. Eine Übersicht über dessen Kennzahlen findet sich in Tabelle 4.1.

Ein Bestandteil der von UA getragenen Avionik sind Sensoren wie DGPS, IMU und Magnetometer. Die Daten dieser Sensoren speisen den Flugsteuerrechner (FCC), der Aufgaben wie automatische Regelung und Pfadpla-



Abbildung 4.25: Unbemannter Helikopter midiARTIS.

nung des UA übernimmt. Ein zweiter Bestandteil der Avionik ist die Hardware, welche die Kommunikation zur Bodenstation ermöglicht. Dies sind drahtlose Kommunikationswege, welche mittels eines seriellen Funkmodems und eines lokalen Netzwerks WLAN aufgebaut werden. Die dritte Komponente der Avionik stellen RC-Komponenten, die zum manuellen Flug und für Eingriffe eines Sicherheitspiloten benötigt werden.

Die Hardware der Umweltwahrnehmung umfasst einen Rechner zur Sensordatenverarbeitung VFC und die jeweiligen Sensoren im hier präsentierten Fall ein monogones LiDAR und eine monokulare Kamera, auf welche in Abschnitt 4.6.1 näher eingegangen wird.

Die Kommunikation zwischen dem FCC und dem VFC erfolgt über ein lokales Netzwerk.

Der ARTIS-Aufbau, welcher Flughardware, Bodenstation und Software umfasst, ist derart konzipiert, dass die Aufgaben eines vollkommen automatischen Flugs bewältigt werden können. Um dies zu ermöglichen, sind an Bord des Helikopters und zwischen diesem und der Bodenstation einige Kommunikationswege notwendig. Einen schematischen Überblick über die verwendeten Datenflüsse und deren Zweck zeigt die Abbildung 4.26. Ausgehend von den Sensoren, DGPS, IMU und Magnetometer, bzw. deren Rohdaten wird innerhalb des FCC die aktuelle Zustandsschätzung mittels einer Sensorfusion [60] ermittelt. Die Informationen des Zustandsschätzers werden von Flugregelungs- und Flugablaufansätzen genutzt. Die resultierenden Kommandos dieser Ansätze werden wiederum von den Aktuatoren zu Stellgrößen umgesetzt um damit Einfluss auf den aktuellen Flugzu-

max. Abflugmasse (MTOW)	12kg
Leermasse	6kg
Masse der Avionik	4,5kg
Masse für Umweltwahrnehmung (Rechner + Sensoren)	1,5kg (Sensoren: \approx 500g)
Rotordurchmesser	1,9m
Antrieb	1,5kW Verbrennungsmotor
Maximale Flugzeit	\approx 20min
Flugsteuerrechner (FCC)	Intel P4 mobile 1,4Ghz, QNX Neutrino OS
Sensordatenverarbeitung (VFC)	Intel Core Duo 1,5GHz, Linux OS

Tabelle 4.1: Kennzahlen der midiARTIS-Konfiguration.

stand zu nehmen. Für einen automatisch geregelten Flug ohne simultane Umweltwahrnehmung ist die alleinige Verwendung des FCC plus präsentierter Sensorik hinreichend. Das setzt allerdings eine hindernisfreie Umwelt voraus.

Wird eine Berücksichtigung der Umwelt für den Flugablauf angestrebt, werden zusätzlich ein VFC sowie Sensorik zur Abtastung der Umwelt benötigt. Zur Herleitung der Umweltinformationen findet vergleichbar zum FCC ein Einsammeln der Sensorrohdaten statt, die anschließend bei Bedarf fusioniert werden. Darauf aufbauend werden die Daten bearbeitet und ausgewertet, um die gewünschten Resultate bereitzustellen. In dem präsentierten Fall ist das gewünschte Resultat eine Hinderniskarte.

Eine korrekte Kombination der Sensordaten und Verortung der Hindernisse im VFC setzt das Wissen über den aktuellen Zustand des Helikopters voraus, welcher von der Sensorfusion des FCC abgefragt wird. Die resultierenden Hindernisinformationen, die Einfluss auf die Flugplanung haben könnten, werden anschließend an diese propagiert.

Für die Einflussnahme an die an Bord befindlichen Komponenten findet zusätzlich eine Kommunikation mit der Bodenstation (GCS) statt. Dies ermöglicht das Abfragen des jeweiligen Status sowie die Anpassung möglicher

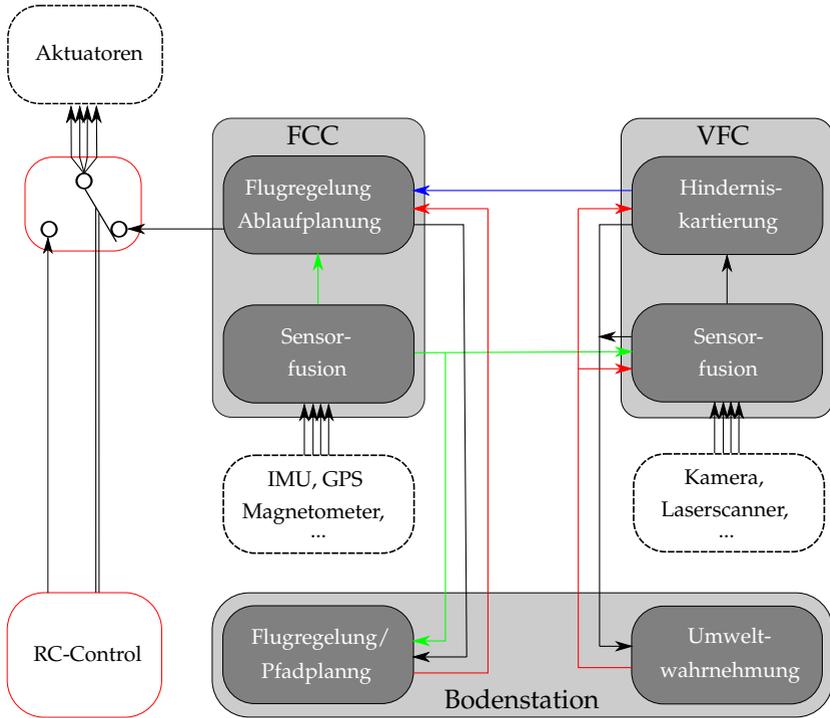


Abbildung 4.26: Schematischer Aufbau der ARTIS-Avionik mit relevanten Datenpfaden.

cher Verhalten. Die Rechner an Bord des Hubschraubers sind zeitlich synchronisiert und die versendeten Daten sind jeweils mit einem Zeitstempel versehen, um eine Zuordnung zu ermöglichen.

Das Sicherheitskonzept des ARTIS-Aufbaus sieht vor, dass der automatische Eingriff auf den Flugzustand des Helikopters durch das an Bord befindliche System jederzeit durch einen Sicherheitspiloten umgangen werden kann. Die Aktuatorstellgrößen des FCC laufen über eine Weiche, die durch den Sicherheitspiloten gesperrt werden kann und den Hubschrauber damit wieder seinen Steuereingaben unterstellt.

Sensoren

Im Rahmen dieser Arbeit wurde die midiARTIS-Konfiguration mit einer monokularen Kamera und einem monogonen LiDAR ausgerüstet. Einen Überblick über deren Eigenschaften zeigt die Tabelle 4.2.

	monokulare Kamera „ImagingSource DMx 21F04“	monogones LiDAR „Hokuyo UTM-30LX“
		
Masse	265g	320g
Sichtfeld	$70^\circ \times 48^\circ$	270°
Bildgröße	$640px \times 480px$	$1081px \times 1px$
Updaterate	30Hz	40Hz
Brennweite	$580px$	-
Max. Reichweite	∞	30m
Spektrum	$\approx 380nm - \approx 750nm$	950nm

Tabelle 4.2: Kennzahlen der Sensorkonfigurationen.

4.6.2 Umgebungskartierung mit Sichtfelderweiterung

Das Ziel der Sichtfelderweiterung für das monogone LiDAR durch den Einsatz einer monokularen Kamera ist, wie der Name schon sagt, die Erweiterung des Sichtkegels der monogonen LiDAR-Sicht im Hinblick auf Reichweite und Sichtfeld. Generalisierbare Werte für die Steigerung der Reichweite und des Sichtfeldes können allerdings nicht genannt werden, da die Sichtfelderweiterung durch die Form und Textur der Hinderniskulisse sowie Beleuchtung des jeweiligen Szenarios bedingt ist.

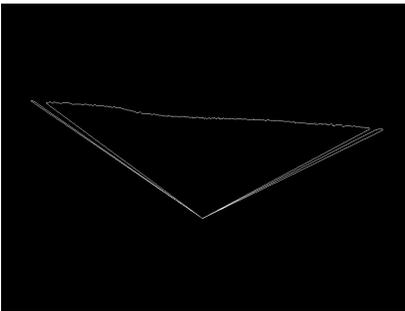
Zur Verdeutlichung der Funktionalität der präsentierten Sichtfelderweiterung wird das vorgestellte Beispielszenario des Überflugs über eine Wiese, auf welcher zwei künstliche Hindernisse aufgestellt sind, herangezogen. Der Überflug dauert $\approx 16,15 \text{ sec}$ und enthält $n_S \approx 646$ Scans.

Im Folgenden werden zuerst anhand eines Beispieldatensets die Einzel-

schritte der Sichtfelderweiterung qualitative vorgestellt und mögliche Schwachpunkte beleuchtet. Die Darstellung der Einzelschritte vermittelt einen ersten Überblick über die Funktionalität des Ansatzes. Die Aussagekraft über die Richtigkeit der gesamten Sichtfelderweiterung ist allerdings begrenzt. Um belastbare Aussagen zur Richtigkeit der generierten Umweltrepräsentation und der Sichtfelderweiterung zu erhalten, folgt eine quantitative Bewertung. Dafür werden zwei OcTrees verglichen, wobei einer die Referenz bzw. Ground Truth und der andere den zu testen Kandidaten darstellt. Das Augenmerk liegt einerseits darauf, dass Hindernisse, welche mit dem Referenz-OcTree gefunden werden ebenfalls in dem OcTree der Sichtfelderweiterung vorhanden sind. Andererseits soll nachgewiesen werden, dass die Sichtfelderweiterung eine umfassendere Sicht auf die Umwelt bereitstellt, als es eine Umweltwahrnehmung erlaubt, welche auf einem monogenen LiDAR basiert.

Einzelschritte der Sichtfelderweiterung

Die Resultate der Einzelschritte werden anhand eines Datensets verdeutlicht, das aus einem Kamerabild besteht dem mehrere Laserscans zu geordnet wurden. Ausgehend von den Rohdaten eines monogenen Lasers (Abbildung 4.27a) und einer monokularen Kamera (Abbildung 4.27b) beginnt die Initialisierung der Sichtfelderweiterung mit der Vorverarbeitung der jeweiligen Daten.



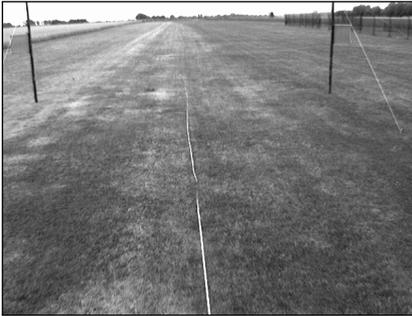
(a) Laserrohdaten, Sicht von oben auf die monogene Ebene.



(b) Kamerarohdaten

Abbildung 4.27: Rohdaten vom monogenen Laserscanner und von der monokularen Kamera.

Vorverarbeitung Die in Abschnitt 4.3.1 vorgestellten Algorithmen der Vorverarbeitung werden auf die Kamerarohdaten des hiesigen Beispiels angewendet. Ausgehend von den Rohdaten erfolgte eine Histogramm-Äqualisation (Abbildung 4.28a), eine Gauß-Pyramide mit zwei Ebenen (Abbildung 4.28b), ein Bilateral-Filter (Abbildung 4.28c) und ein Canny-Operator (Abbildung 4.28d). Das Resultat der Histogramm-Äqualisation zeigt ge-



(a) Resultat nach Histogramm-Äqualisation



(b) Resultat nach Gauß-Pyramide



(c) Resultat nach Bilateral-Filter



(d) Resultat nach Canny-Operator

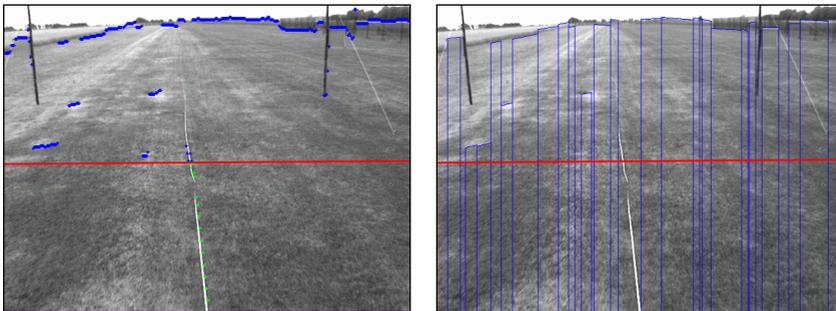
Abbildung 4.28: Vorverarbeitung der Kameradaten.

genüber dem Rohbild ein Steigerung des Dynamikumfangs, was zu dem gewünschten Hervorheben von Intensitätsdiskontinuitäten führt. Die Unschärfe in den Resultaten der Gauß-Pyramide und dem Bilateral-Filter sind zum einen auf die Halbierung der Auflösung gegenüber dem Ausgangsbild zurückzuführen. Zum anderen soll die Filterung geringe Intensitätsdiskontinuitäten unterdrücken, um Rauschen in homogenen Flächen zu reduzieren. Diese Tiefpassfilterung macht sich ebenfalls als Weichzeichnung des Bildes bemerkbar. Das Resultat des Canny-Operators hebt so-

wohl die Kanten zu den beiden aufgestellten Hindernissen als auch den Übergang zwischen Boden und Himmel sehr gut hervor. Es ist aber auch ersichtlich, dass Kanten, welche nicht auf räumlichen Kanten beruhen, wie z.B. das Maßband, das sich in der unteren Mitte des Bildes befindet, ebenfalls hervorgehoben werden und somit zu einer Einschränkung der maximal möglichen Sichtfelderweiterung führen.

Auf die Vorverarbeitung der LiDAR-Daten wird nicht eingegangen, da es sich um Transformationen handelt, die zu unterschiedlichen Zeitpunkten mit unterschiedlichen Parametern durchgeführt werden und sich damit schlecht visuell aufbereiten ließen. Der letzte Schritt der Vorverarbeitung der Laserdaten ist die Projektion des zuletzt generierten Scans auf die Bildebene. Die projizierten Lasermessungen sind in Abbildung 4.29a rot markiert.

Segmentierung Auf die Vorverarbeitung, der Sensordaten folgt die Segmentierung des Kamerabildes ausgehend von den projizierten Laserdaten. Das Resultat des ersten Schritts, die Identifikation der Kantenpixel, zeigt



(a) Identifikation von Kantenpixeln

(b) Konstruktion von Segmenten

Abbildung 4.29: Segmentierung der Bildebene.

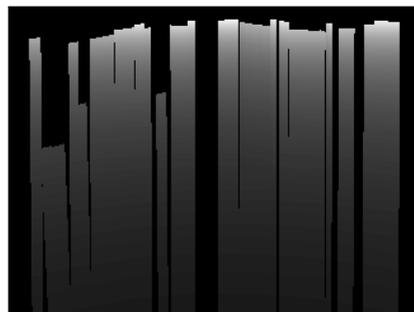
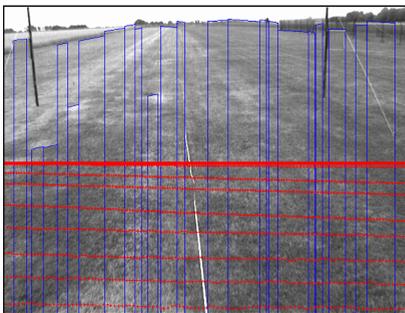
Abbildung 4.29a. Zur besseren Visualisierung, welche Kantenpixel ausgewählt wurden, wurde als Hintergrund nicht das Canny-Bild, sondern das Originalbild der Kamera verwendet. Die auf die Bildebene projizierten Lasermessungen entsprechen aufgrund der getroffenen Annahmen von planaren Objektoberflächen der Projektion der Schnittgeraden \mathbf{g}_{OA} zwischen der Scanebene und der betrachteten Objektoberfläche. Die Projektion von \mathbf{g}_{OA} verläuft nahezu parallel zur x -Achse des Bildes. Daraus resultiert eine Suche nach möglichen Segmentkanten, parallel zur y -Achse des Bildes. Die gefundenen Kanten sind im Bild blau markiert, wenn sie sich oberhalb der

Laserprojektionen befinden, und grün, wenn sie unterhalb der Laserprojektionen liegen. Ein Vergleich zwischen den markierten Kanten in 4.29a und denen in Abbildung 4.28d zeigt, dass immer die nächstliegende Kante zur korrespondierenden Laserprojektion markiert wurde. Die identifizierten Kantenpixel für den Bereich unterhalb der projizierten Lasermessungen liegen zum größten Teil auf dem Bildrand.

Das Ergebnis der Segmentierung zeigt Abbildung 4.29b. Die konstruierten Segmente haben die Form eines Trapezes, wobei die Grundkanten parallel zur Suchrichtung des vorhergehenden Schritts verlaufen und die Schenkel auf den gefundenen Kantenpixeln beruhen. Die Segmente sind in der Abbildung blau umrahmt.

3D Rekonstruktion Aufbauend auf der Segmentierung erfolgt die Übertragung der Segmente in den Objektraum. Wie erläutert, genügt die alleinige Verwendung des zuletzt generierten Scans nicht, um die Pose einer planaren Oberfläche im Objektraum eindeutig zu bestimmen. Es werden weitere Lasermessungen, welche auf älteren Scans basieren, auf die Bildebene projiziert und den Segmenten zugeordnet, um jedem Segment eine linear unabhängige Punktmenge bereitzustellen. Abbildung 4.30a zeigt die Projektionen von acht weiteren Scans auf die Bildebene. Die Verteilung über die Bildebene erfolgt durch die Bewegung des UA während der Aufzeichnung der einzelnen Scans.

Mittels der zugeordneten Projektionen bzw. deren Objektkoordinaten wer-



(a) Zuordnung mehrerer Scans zu einem Kamerabild (b) Tiefenbild der generierten Objektberflächenebenen

Abbildung 4.30: Resultat der kooperativen Fusion.

den dreidimensionalen Regressionsebenen berechnet, welche durch die ebenfalls in den Objektraum projizierten Segmentgrenzen begrenzt werden.

Abbildung 4.30b zeigt die ermittelten dreidimensionalen Objektoberflächen als Tiefenbild. Der Graufärbung innerhalb des Bildes zeigt die kartesische Entfernung zwischen dem Kameraursprung und dem auf die Bildebene projizierten Objektpunkt. Dunkle Töne zeigen eine geringe und helle eine größere Entfernung zum Kameraursprung an. Die dargestellten Objektoberflächen beschreiben in dem vorgestellten Beispiel den Erdboden. Aus diesem Grund werden sie vom unteren Bildrand nach oben heller, was einen Zuwachs der Entfernung beschreibt. Die Entfernungsdaten innerhalb des Tiefenbildes liegen nicht mehr als zusammenhängende Ebenen vor. Sie wurden für die Erstellung des Tiefenbildes diskretisiert und können der konkurrierenden Fusion als einzelne Approximationen zugeführt werden. Die resultierenden Approximationen werden anschließend in einen OcTree überführt. Einen beispielhaften OcTree für den Zeitschritt eines Kamerabildes zeigt Abbildung 4.31. Die Einfärbung des Oc-

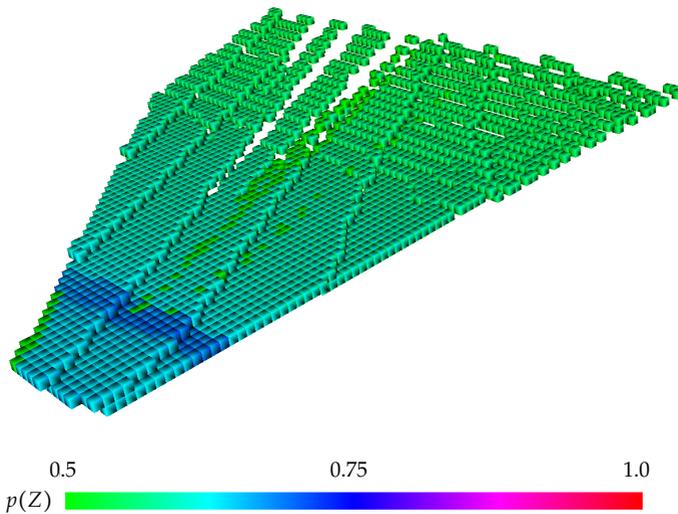


Abbildung 4.31: OcTree, wobei die Färbung der Voxel deren Belegungswahrscheinlichkeit $p(Z)$ widerspiegelt.

Trees ist auf das verwendete Sensormodell zurückzuführen und spiegelt die Belegungswahrscheinlichkeit $p(Z)$ wider. Zu erkennen ist der blaue Bereich, für welchen Lasermessungen vorliegen und damit die Belegungswahrscheinlichkeit der Approximation mit $p(Z) \approx 0.8$ in etwa der von Laserrohdaten entspricht. Des Weiteren lässt sich mit einer zunehmenden Entfernung vom Kameraursprung bzw. den Lasermessungen (Abschnitt:

4.5.1) eine Reduktion der Belegungswahrscheinlichkeit erkennen. Die kontinuierliche Reduktion von $p(Z)$ wird von Bereichen durchzogen, welche noch eine geringe Aussagekraft besitzen. Diese Bereiche sind auf die erstellten Regressionsebenen zurückzuführen, denen ein verringertes Vertrauen ausgesprochen wurde, weil sie die jeweils zugrunde liegende Punktmenge nicht hinreichend repräsentierten (siehe Abschnitt 4.5.1).

Diskussion Für die dargebrachten Beispieldaten von Laserscanner und Kamera zeigen die Abbildungen, dass der Ansatz bzw. das Vorgehen die angestrebte Vergrößerung des Sichtfeldes herbeiführt. Wie die Qualität der entstehenden Approximationen ist, kann allein auf Basis der Visualisierung nicht festgestellt werden. Auf diesen Punkt wird in Abschnitt 4.6.2 eingegangen.

Es lassen sich allerdings auf Basis der gezeigten Abbildungen Schwächen des präsentierten Ansatzes ausmachen, auf welche im Folgenden eingegangen wird.

Im Kamerabild besteht die Möglichkeit, dass Pixel, die im Dreidimensionalen unterschiedliche Objekte abbilden, auf der Bildebene benachbart und mit einer vergleichbaren Intensität ausgestattet sind. Die Folge ist, dass die verwendete Kantendetektion die Übergänge zwischen den Objekten nicht korrekt identifiziert. Resultierend werden Regressionsebenen erzeugt, deren Grenzen keine reale Grundlage besitzen. Beispiel für diese Problema-

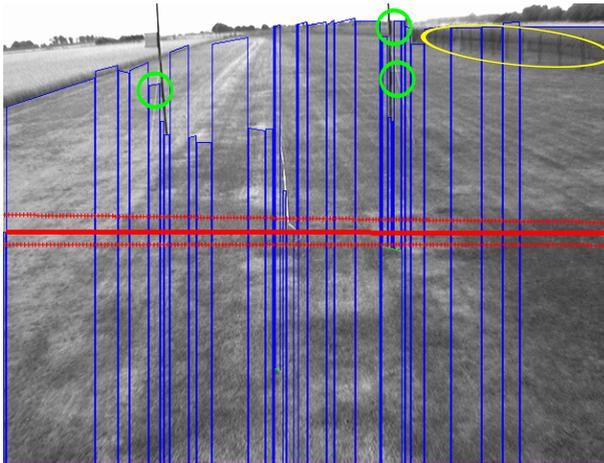


Abbildung 4.32: Detektionsfehler bei der Bestimmung der Segmentgrenzen, als Resultat einer nicht korrekten Kantendetektion.

tik sind die Flaggen an den Pfosten oder der Zaun in der oberen rechten Ecke etwa in Abbildung 4.29b. Zur Verdeutlichung zeigt Abbildung 4.32 die fälschlich zugeordneten Flaggen mit grünen Kreisen und den Zaun mit einer gelben Ellipse. Die Pfosten, die als Hindernisse in dem Szenario platziert wurden, werden aufgrund der starken Intensitätsdiskontinuität gegenüber dem Boden wie erwünscht erkannt. Die an den Pfosten angebrachten Flaggen werden von der Sonne durchstrahlt, was ihre Intensitäten vergleichbar zu denen des Bodens erscheinen lässt. Beim Zaun wird davon ausgegangen, dass die horizontalen Strukturen des Zaunes zu fein sind und vom Canny-Detektor als Rauschen bzw. schwache Kanten identifiziert werden und dadurch verworfen werden. Erst der Übergang von der Flagge bzw. dem Zaun zum Himmel stellt eine hinreichende Schwelle für den Canny dar.

Diese Fehler lassen darauf schließen, dass feine Strukturen, wie etwa Stromleitungen oder Drahtschlaufenzäune eine starke Fehlerquelle für den Ansatz darstellen.

Ein weiterer Schwachpunkt, den Abbildung 4.32 andeutet, ist, dass die Regressionsebenen pro Segment jeweils auf einer kleinen Basis an Lasermessungen beruht. Im dargestellten Fall handelt es sich um drei Scans. Die Problematik wurde frühzeitig erkannt und im Abschnitt 4.4.3 durch die Kombination mehrerer Ebenen behandelt und in Abschnitt 4.5.1 als ein Ausgangspunkt für die Entwicklung des Sensormodells präsentiert. Die resultierende Reduktion der Aussagekraft ist anhand der Belegungswahrscheinlichkeit $p(Z)$ in Abbildung 4.31 verdeutlicht.

Quantitative Validierung

Referenzdaten Um quantifizierbare Aussagen über die Übereinstimmung des aus der Sichtfelderweiterung resultierenden OcTree und der realen Umwelt zu erreichen, wird eine Ground Truth benötigt, welche das Aussehen der realen Umwelt widerspiegelt. Als Referenz werden in der hier präsentierten Validierung rohe LiDAR-Messungen verwendet, welche für die betrachtete Umwelt vorliegen. Mittels dieser Messungen wird ein OcTree OT_T in Abbildung 4.33 erzeugt, welcher alle zur Verfügung stehenden $i_S = 646$ Scans enthält.

Ausgangspunkt für die Entscheidung, als Ground Truth die LiDAR-Rohdaten zu verwenden, ist, dass es sich bei dem hier präsentierten Ansatz um eine Erweiterung der monogonen LiDAR-Sicht handelt, deren gewünschtes Ziel es ist, mindestens die Hindernisse zu detektieren, welche mittels der Laserrohdaten erkannt werden und wenn es die Sichtbedingungen erlauben, Approximationen von Hindernisstrukturen zu präzisieren, die

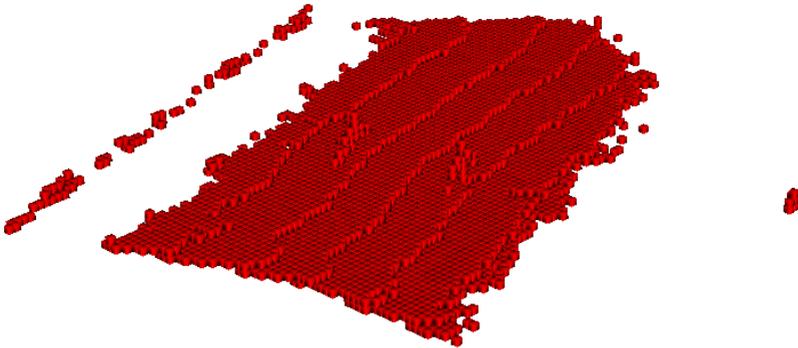


Abbildung 4.33: OcTree OT_T des Beispielszenarios, erstellt allein mit LiDAR-Daten des monogenen Scanners.

über das Sichtfeld des LiDAR hinausgehen.

Da erwartet wird, dass die Sichtfelderweiterung eine größere Anzahl Hindernisstrukturen erkennt, als in der Ground Truth enthalten sind, wird das Beispielszenario dahingehend angepasst. Bei dem gegebenen Szenario und Flugpfad liegen nur drei Hindernisse - die beiden Torpfosten und der Erdboden - vor. Die beiden Torpfosten werden mittels des LiDAR nahezu vollständig erkannt. Der Boden reicht über das mit dem LiDAR betrachtete Gebiet hinaus. Aus diesem Grund wird davon ausgegangen, dass mit der Fusion auch Bodenbereiche erkannt werden, welche in der Ground Truth nicht vorliegen.

Basierend auf dem Vorwissen über das beflogene Gebiet und eine Bewertung der Bodenmessungen im Laserdatensatz, wird der Boden als eine Ebene angenommen. Für die Ermittlung der Ebenenparameter wurden Messungen der Ground Truth, die nicht den Boden repräsentieren, manuell entfernt und die verbliebenen Messungen zu einer Regressionsebene E_G vereinigt. Abbildung 4.34 zeigt den OT_T , der um die Voxel erweitert wurde, welche von E_G geschnitten werden. Der resultierende OcTree OT_S wird im Weiteren als Signal bezeichnet. Zur Generierung von OT_T werden 274036 Messungen verwendet. Als Basis für E_G werden 99,4% dieser Messungen genutzt. Angaben dazu, wie E_G diese Punktmenge abbildet, zeigt Abbildung 4.35. Durch die manuelle Auswahl wird eine Punktmenge bereitgestellt, die durch E_G sehr gut repräsentiert wird. Der maximale orthogonale Abstand, den ein Punkt zur Ebene aufweist sind $0,55\text{ m}$, was etwa der Auflösung des hier verwendeten OcTree entspricht. Ein Median von maximal

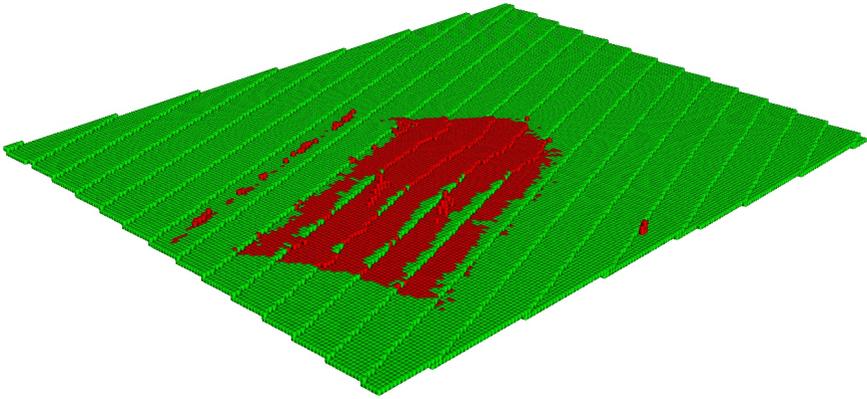


Abbildung 4.34: OcTree OT_T des Beispielszenarios erweitert mit Regressionsebene E_G .

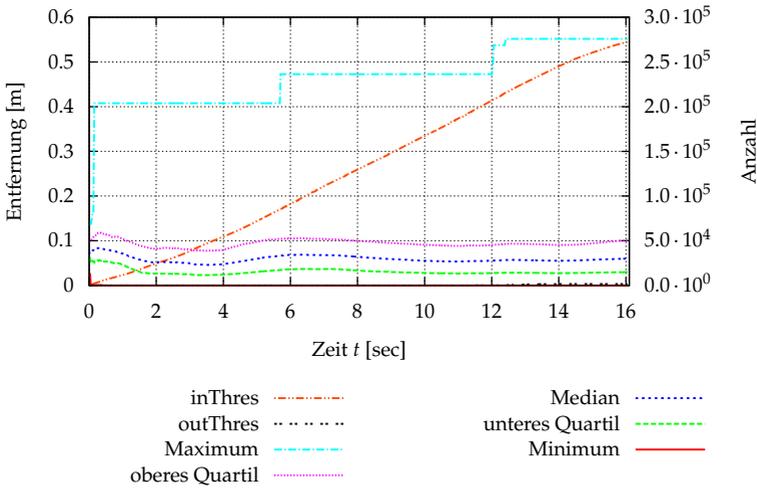


Abbildung 4.35: Verifikation der Regressionsebene E_G gegenüber Laserrohdaten.

0,08 m bzw. ein Wert von ca. 0.1 m für das obere Quantil zeigen allerdings, dass die größte Anzahl der Abweichungen unter der Auflösung des OcTree liegt. Eine Übertragung von E_G in den OcTree lässt damit eine hinreichende Repräsentation des Bodens erwarten. Der OcTree OT_S bildet die erweiterte Ground Truth, gegen welche der OcTree OT_F der Sichtfelderweiterung verglichen wird.

Signalentdeckungstheorie Dieser Vergleich wird unter Anwendung der Signalentdeckungstheorie [34] umgesetzt. Die Signalentdeckungstheorie analysiert die Detektion eines zu identifizierenden Signals und stellt eine Maßzahl für die Qualität der Detektion bereit.

Das zu detektierende Signal wird im präsentierten Fall durch die belegten Zellen von OT_S bzw. OT_T bereitgestellt. Die Detektion erfolgt mittels eines OcTree, welcher die zu testende Umwetrepräsentation enthält. Für die Eruiierung werden die belegten Voxel im zu bewertenden Baum (OT_L bzw. OT_F), deren Position im Baum eindeutig ist, mit den entsprechenden Voxeln im Signal-Baum (OT_T bzw. OT_S) verglichen. Ist der betrachtete Voxel ebenfalls belegt, wurde das Signal korrekt identifiziert. Um eine Einordnung der Detektion zu erreichen, wird eine binäre Dekomposition des Problems vorgenommen. Tabelle 4.3 verdeutlicht die verwendete Kontingenztafel. Voraussetzungen für dieses Vorgehen sind, dass die Wurzeln

		Signal	
		Zelle m in OT_S belegt	Zelle m in OT_S frei
Detektion	erkannt	belegte m wird korrekt erkannt	freie m wird als belegt erkannt
	nicht erkannt	belegte m wird als frei erkannt	freie m wird korrekt erkannt

Tabelle 4.3: Kontingenztafel der verwendeten Signalentdeckungstheorie.

der verglichenen Bäume eine identische Pose aufweisen und dass die Auflösung der Bäume identisch ist. Als Referenz für die Funktionsfähigkeit dieses Vorgehens wird der OcTree OT_T mit dem inkrementell erstellten OcTree $OT_{L,1:i_S}$ nach jedem Inkrement i_S sukzessive verglichen. $OT_{L,1:i_S}$ beruht allein auf Lasermessungen und dient als Kontrollinstanz. Der OcTree OT_T besitzt eine Tiefe von $h = 12$ und die Messungen wurden mit dem Ansatz $C(4)$ (Abschnitt 3.2.5) integriert.

Bei der Gegenüberstellung von OT_T und $OT_{L,1:i_S}$ wird erwartet, dass nach

der Integration der Inkremente $i_s = 1 - 646$ beide Bäume identisch sind. Nach einem Inkrement $i_s \neq 646$ besteht die Möglichkeit, dass Zellen in $OT_{L,1:i_s}$ eine Belegung aufweisen, welche in OT_T gegensätzlich deklariert ist. Dies tritt durch die probabilistische Bewertung der Zellen auf und kann durch eine Anpassung der entsprechenden Zelle durch nachfolgende Inkremente entstehen, etwa wenn eine Zelle zu unterschiedlichen Zeitpunkten aus unterschiedlichen Blickwinkeln betrachtet wird.

Das Resultat der Gegenüberstellung von $OT_{L,1:i_s}$ und OT_T zeigt Abbildung 4.36. Der Graph $n_{Tr} = 4831$ zeigt die Anzahl der belegten Voxel in OT_T .

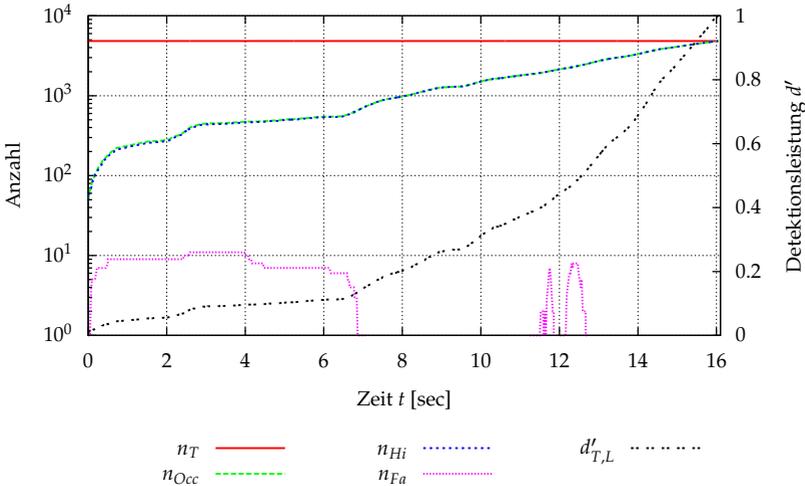


Abbildung 4.36: Ausgangssituation allein mit Laserdaten.

n_{Occ} verdeutlicht die Anzahl der belegten Voxel in $OT_{L,1:i_s}$ für das Inkrement i_s . Ausgehend von der Anzahl der belegten Zellen n_{Occ} in $OT_{L,1:i_s}$ sind n_{Hi} Voxel ebenfalls in OT_T belegt. Die Differenz zwischen n_{Occ} und n_{Hi} spiegelt eine Abweichung der beiden OcTrees wider und wird mittels n_{Fa} ausgedrückt. Die Detektionsleistung

$$d'_T = \frac{n_{Hi}}{n_{Tr}} - \frac{n_{Fa}}{n - n_{Tr}} \tag{4.16}$$

zeigt an, wie gut $OT_{L,1:i_s}$ zum Zeitpunkt $t \approx i_s @ 40\text{Hz}$ dem OcTree OT_T entspricht, wobei n die theoretisch im Baum vorhandene Anzahl von Blättern ist.

Die Bäume $OT_{L,1:i_s}$ und OT_T basieren auf einer identischen Datenbasis, den

reinen monogonen LiDAR-Daten. Es war daher zu erwarten, dass nach der Integration des letzten Inkrements $i_S = 646$ beide Bäume identisch sind. Dieses erwartete Resultat tritt ein, und wird durch $d'_T = 1$ in Abbildung 4.36 verdeutlicht.

Aufbauend auf dem Vergleich von $OT_{L,1:i_S}$ und OT_T wird davon ausgegangen, dass die Signalentdeckungstheorie geeignet ist, zwei OcTrees miteinander zu vergleichen, um die Richtigkeit des OcTree der Sichtfelderweiterung $OT_{F,1:i_S}$ gegenüber der Ground Truth OT_S zu bewerten.

Die Resultate der Gegenüberstellung von $OT_{F,1:i_S}$ und OT_S zeigen die Abbildungen 4.37 und 4.38. Abbildung 4.37 visualisiert mit n_{Occ} die Gesamtanzahl an belegten Voxeln in $OT_{F,1:i_S}$ nach der Integration des Scans i_S . Die Teilmenge von n_{Occ} , welche mit der Ground Truth in OT_S übereinstimmt, zeigt n_{Hi} . n_{Pl} enthält die Anzahl der Voxel, welche mit der künstlichen Ebene E_G in OT_S übereinstimmen und n_{Fa} ist die Anzahl der falsch zugeordneten Voxel zwischen $OT_{F,1:i_S}$ und OT_S . Zur Einordnung wurde noch die Anzahl der belegten Voxel in OT_T mit n_T und die Anzahl der Voxel in OT_S mit n_S abgebildet.

Im Vergleich zwischen Abbildung 4.36 und 4.37 fällt auf, dass von Anfang

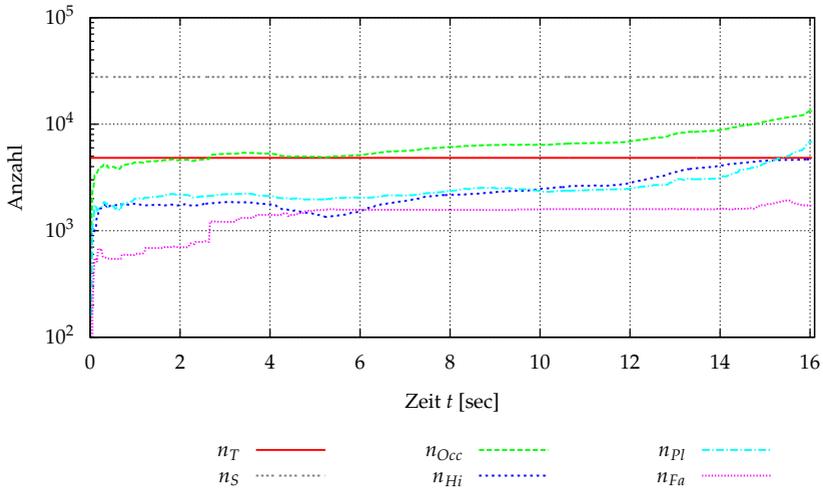


Abbildung 4.37: Inkrementelle Dekomposition der Signaldetektion für den OcTree $OT_{F,1:i_S}$ gegenüber dem Signal OT_S .

an die jeweiligen Werte von n_{Occ} , n_{Hi} und n_{Fa} pro Zeitschritt in Abbildung 4.37 über denen in Abbildung 4.36 liegen. Die größere Anzahl von n_{Occ} spricht aufgrund der identischen Auflösung des OcTree für eine umfang-

reichere Betrachtung der Umwelt, bzw. dafür, dass pro Scan/Zeitschritt eine größere Anzahl von Messungen für die Sichtfelderweiterung vorliegt als für die reine Vermessung mit LiDAR.

Neben Aussagen über den gesteigerten Umfang der Umweltbetrachtung lassen sich auch Aussagen über die Qualität der Umweltbetrachtungen treffen. Der Wert n_{Fa} steigt bis zum Zeitpunkt $t \approx 2,8s$ sprunghaft an. Ab diesem Punkt ist der Verlauf des Graphen eher konstant. Dem gegenüber steigt n_{Occ} über den gesamten Verlauf der 646 Scans. Diese Beobachtung wirft zwei Interpretationen auf. Einerseits, es wird zu Beginn der Umweltbeobachtung eine große Anzahl von Falschmessungen in den OcTree integriert und nach einer Phase, in der eine hinreichende Datenbasis für die LiDAR für die Kombination mit einem Kamerabild aufgebaut wurde, werden nur korrekte Messungen in den Baum hinzugefügt. Da die Graphen n_{Hi} und n_{Pl} , welche eine korrekte Signaldetektion beschreiben, nach dieser Phase keine sprunghaften Anstiege aufweisen, ist diese Interpretation aber unwahrscheinlich.

Auf der anderen Seite die wahrscheinlichere Annahme, dass Messungen, welche in den Baum integriert werden und falsch gegenüber OT_5 sind, auf Basis des probabilistischen Charakters des OcTree teilweise durch nachgelagerte Scans korrigiert werden. Gleichzeitig erfolgt pro Scan die Integration neuer Falschmessungen, wobei sich die Anzahl der korrigierten und der neu hinzugefügten Falschmessungen aufgrund der konstanten Sichtfeldgröße und des stetigen Flugpfades etwa die Waage hält. Diese Interpretation unterstützt die Erwartung an die Sichtfelderweiterungen, dass Oberflächen, welche teilweise auf falschen Prognosen beruhen, über die Zeit korrigiert und deren wahre Formen herausgearbeitet werden.

Eine Korrektur von Voxeln ist nur möglich, wenn die durch den Voxel beschriebene Umwelt mehrfach abgetastet wird. Dies setzt voraus, dass der Sichtkegel der Sensorfusion mehrfach die entsprechende Region enthält. Eine sprunghafte Änderung der Fluglage, welche eine kontinuierliche Änderung des OcTree verhindert, kann dazu führen, dass in den OcTree eingetragene Messungen nicht korrigiert werden und folglich das Level der Falschmessungen n_{Fa} im OcTree wiederum ansteigt. Ein weiterer Punkt, der sich aus Abbildung 4.37 entnehmen lässt, ist dass n_{Hi} schlussendlich gegen n_{Tr} konvertiert, was eine vollständige Detektion aller Hindernisse verdeutlicht, welche auch mit der reinen Laserdetektion erfasst wurden. Dieses Ergebnis bestätigt die Erwartung, dass mit der Sichtfelderweiterung mindestens die Hindernisse erkannt werden, welche mit der reinen Laserdetektion erkannt werden. Zusätzlich zu der Beobachtung, dass die dargestellten Größen n_{Occ} , n_{Hi} und n_{Fa} pro Zeitschritt in Abbildung 4.37 über denen in Abbildung 4.36 liegen, zeigt die Größe n_{plane} ebenfalls einen

Zuwachs des Sichtfeldes an. Die Ebene E_G ist in OT_T nicht enthalten und es liegen folglich auch keine Signaldetektionen für die Voxel, die Ebenen repräsentierenden, vor. Jede Detektion eines dieser Voxel durch die Signaldetektion verdeutlicht eine Erweiterung des Sichtfeldes und damit eine umfassende Detektion der Hindernisse, die über die reine Detektion mit einem LiDAR hinausgeht.

Abbildung 4.38 verdeutlicht noch einmal die in Abbildung 4.37 visuali-

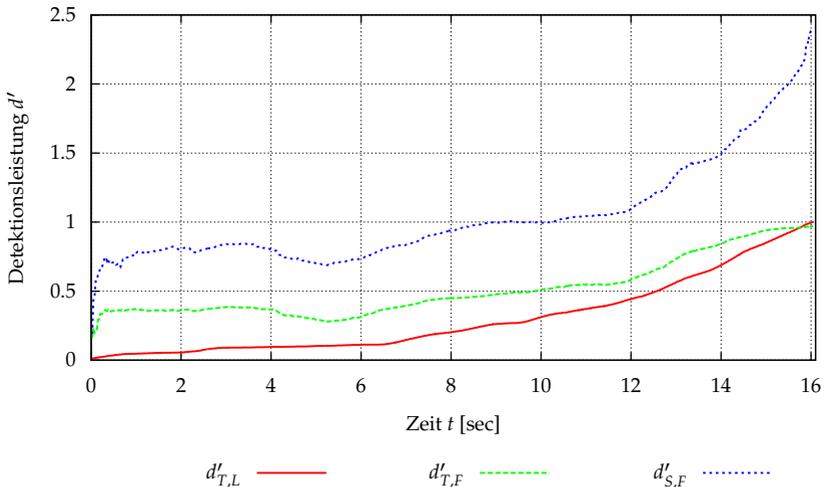


Abbildung 4.38: Detektionsleistung basierend auf der Gegenüberstellung von $OT_{L,1:i_S}$ und OT_T sowie $OT_{F,1:i_S}$ und OT_S .

sierten Resultate. Der Graph $d'_{T,L}$ ist identisch zu der Darstellung in Abbildung 4.36 und lässt sich mittels 4.16 herleiten.

Aufbauend auf Gleichung 4.16 wurde ebenfalls der Graph $d'_{T,F}$ ermittelt. Im Vergleich zwischen $d'_{T,L}$ und $d'_{T,F}$ ist zu erkennen, dass $d'_{T,F}$ von Beginn des Überfluges an eine höhere Detektionsleistung aufweist als $d'_{T,L}$, was bedeutet, dass die Sicht auf die Hinderniskulisse mit der Sichtfelderweiterung pro Scan umfangreicher ist als mit der reinen LiDAR-Sicht. Im Verlauf konvertieren die beiden Graphen gegen 1, was eine 100% Detektion der vorhandenen Hinderniskulisse entspricht. Das Konvertieren von $d'_{T,L}$ gegen $d'_{T,F}$ bedeutet nicht, dass sich die reine LiDAR-Sicht der erweiterten Sensorsicht in seiner Detektionsleistung annähert. Der Grund der Annäherung liegt vielmehr darin, dass keine weiteren Hindernisse vorhanden sind, welche durch die erweiterte Sensorsicht detektiert werden können.

Diese führt dazu, dass beide Kurven abflachen und gegen 100% konvertieren, was zu der dargestellten Annäherung beider Kurven führt.

Einen Eindruck, wie die Detektionsleistung der Sichtfelderweiterung über die erweiterte Hinderniskulisse in OT_S arbeitet, vermittelt der Graph $d'_{S,F}$. Für diesen Graph werden nicht ausschließlich die Werte von n_{Hit} für eine positive Signalidentifikation berücksichtigt, sondern eine Kombination von n_{Hi} und n_{Plane} . Die Gegenüberstellung von $d'_{T,L}$ und $d'_{S,F}$ zeigt zu $d'_{T,F}$, dass nicht nur die Hindernisse in der Kamerasicht frühzeitiger und vollständig mit der Sichtfelderweiterung erkannt werden, sondern dass zusätzlich Hindernisse erkannt werden, welche nicht, auch nicht zu späteren Zeitpunkten, in der LiDAR-Sicht vorliegen.

Schlussendlich lässt sich festhalten, dass für dieses Szenario - mit seiner Hinderniskulisse, dem Flugpfad und den eingesetzten Sensoren - mit der Sichtfelderweiterung eine umfassendere Sicht auf die das UA, umgebende Umwelt ermöglicht wird. Diese umfassendere Sicht lässt sich an der gesteigerten Detektionsleistung $d'_{S,F}$ gegenüber $d'_{T,F}$ festmachen. Dabei handelt es sich im präsentierten Fall zum Zeitpunkt $t \approx 16.23$ um eine Steigerung auf 240%. Wie angedeutet, steht dieser Wert ausschließlich für das hier beschriebene Szenario. Des Weiteren sollte Abstand von einer prozentualen Aussage genommen werden. Das Sichtfeld der präsentierten Erweiterung ist maximal auf das Sichtfeld der Kamera begrenzt, was dafür spricht, dass sich der Mehrwert der Erweiterung maximal auf dieses Sichtfeld beziehen kann. Folglich kann sich die Steigerung prozentual nur auf einen Scan beziehen und nicht auf ein gesamtes Szenario. Bei einem Szenario muss vielmehr von einem additiven Offset ausgegangen werden.

Fazit

Die gestellten Anforderungen, dass mit der Sichtfelderweiterung mindestens alle Hindernisse detektiert werden, welche mit der alleinigen Verwendung des LiDAR wahrgenommen werden, konnte gezeigt werden. Darüber hinaus konnte eine Steigerung der gesichteten Objekte für Sichtfelderweiterung um 140% in dem präsentierten Szenario erreicht werden. Ebenso wurde gezeigt, dass durch den probabilistischen Aufbau der Hindernisrepräsentation eine Anpassung der getroffenen Approximationen über die Zeit stattfindet und dass damit ein Herausformen der wahren Hindernisgestalt möglich ist.

Es wurden allerdings auch Schwächen der Sichtfelderweiterung sichtbar. Negativ fiel etwa die Kombination von Oberflächen in der Bildebene auf, welche in der Realität keine Verbindung aufwiesen. Dieser Fehler konnte auf eine nicht hinreichende Intensitätsdiskontinuität zwischen den jeweili-

gen Oberflächen zurückgeführt werden. Das Eintreten dieser Fehlerquelle kann nicht verhindert werden und kann bei wechselnden Beleuchtungsverhältnissen jederzeit auch in anderen Versuchsaufbauten auftreten. Es wird allerdings davon ausgegangen, dass der dargestellte Fehler geringe bis keine Auswirkungen auf eine korrekte Hindernisdetektion und -repräsentation hat, da bei einer näheren Betrachtung der Hindernisse, das heißt, wenn ein geringerer Abstand zwischen Sensorquelle und Objekt vorliegt, ebenfalls Messungen des LiDAR auf den entsprechenden Oberflächen vorliegen werden und somit eine korrekte Detektion der Hinderisoberflächen gegeben ist.

Kapitel 5

Zusammenfassung und Ausblick

Die vorliegende Arbeit widmet sich der dreidimensionalen Detektion von Hindernissen für kleine unbemannte Luftfahrzeuge mittels einer Sensordatenfusion von monogonen LiDAR- und monokularen Kameradaten. Der vorgestellte Ansatz legt den Fokus auf die Detektion von statischen und dynamischen Objekten während des Fluges und die unverzügliche Bereitstellung der gewonnenen Informationen.

Bei der Auslegung des vorgelegten Ansatzes wird sich auf kleine unbemannte Luftfahrzeuge (UA) konzentriert. Der Betrieb von UA und deren Weiterentwicklung stehen im Fokus von Forschung und Anwendung, da sie die Erschließung neuer Anwendungsfelder und Aufgabengebiete versprechen. Der bodennahe Betrieb von UA in hindernisreichen Umgebungen erfolgt aktuell zum größten Teil noch manuell. Ein automatischer Betrieb, besonders außerhalb des Einflussbereiches eines Operators, erfordert eine Automatisierung der Hindernisdetektion und die unverzügliche Bereitstellung der gewonnenen Daten während des Fluges an eine vorhandene Hindernisvermeidung.

Kleine UA-Systeme können aufgrund ihrer Baugröße und maximaler Nutzlast nur begrenzt Sensoren zur Hindernisdetektion transportieren. Oft stehen nur zweidimensional oder dreidimensional messende Sensoren mit einem stark begrenzten Sichtfeld zur Verfügung. Die begrenzten Fähigkeiten, welche die eingesetzten Sensoren bereitstellen, führen oft zu einer eingeschränkten Einsetzbarkeit des UA. Die Konzipierung einer Hindernisdetektion ist zwangsläufig mit der Frage verbunden, welche Sensoren und welche Techniken eingesetzt werden, um dem UA das dreidimensionale Sehen in hinreichender Qualität, Reichweite und hinreichendem Sichtfeld zu ermöglichen.

Die beschränkten Fähigkeiten der Sensoren machen den Einsatz eines einzelnen Sensors oft nicht hinreichend. Die Fusion von Daten mehrerer Sensoren mit unterschiedlichen Sichtfeldern oder Messmethoden ist daher oft ein Mittel der Wahl, um eine hinreichende Hinderniswahrnehmung für

UA aufzubauen.

Die Kombination von Daten unterschiedlicher Sensoren setzt oftmals eine Zuordnung von korrespondierenden Daten voraus. Zur Herleitung dieser Zuordnung wird eine Kalibrierung der extrinsischen Parameter zwischen Kamera und LiDAR vorgestellt, welche neben den vom LiDAR bereitgestellten Entfernungsmessungen ebenfalls deren Intensitätswerte nutzt, um Irritationen wie „Kantentreffer“ bei der Detektion eines Kalibrierobjekts zu identifizieren. Zusätzlich ermöglicht die Nutzung der Intensitätsmessungen eine automatische Detektion des Kalibrierobjekts in den LiDAR-Daten, was wiederum die Verarbeitung größerer Datenmengen ohne manuellen Eingriff erlaubt. In Simulationen und Experimenten konnte gezeigt werden, dass die Verwendung einer Kombination aus Entfernungs- und Intensitätsmessungen eine genauere Bestimmung der extrinsischen Parameter erlaubt, als es eine rein auf Entfernungsmessungen basierende Detektion vermag.

Einzelne Sensordaten können mit Messrauschen belastet sein oder gar Falsch-aussagen enthalten. Um eine gesicherte Aussage über eine betrachtete Messgröße zu erhalten ist eine probabilistische Kombination mehrerer unabhängiger Messung eine adäquate Methode. Im Bereich der Hindernisdetektion bzw. -repräsentation haben sich zu diesem Zweck diskrete, metrische Belegtheitsgitter bewährt. Bei der Überführung der Sensordaten aus der kontinuierlichen Umwelt kann es aufgrund der Skalierung der diskretisierten Repräsentation allerdings zu widersprüchlichen Aussagen pro Diskret kommen, ob ein Diskret durch ein Hindernis belegt ist oder nicht. Dieser Widerspruch kann zu einer Situation führen, in der Bereiche der Repräsentation als frei von Hindernissen deklariert werden, obwohl in der realen Welt für diese Bereiche Hindernisse vorliegen. Diese Diskrepanz kann zu Kollisionen des UA mit Hindernissen und damit zum Verlust des UA führen.

Zur Behebung des Problems, der ungewollten Freizeichnung von statischen Hindernissen, wurden drei Ansätze präsentiert, welche die auf die Diskretisierung zurückzuführende Freizeichnung verdecken. Eine Gegenüberstellung der drei Ansätze stellt die Vorteile bzgl. Rechen- und Speicheraufwand eines Ansatzes heraus, der die einzelnen Diskrete der Hindernisrepräsentation nicht, wie üblich, als unabhängig, sondern entlang des Messstrahls als bedingt verknüpft behandelt.

Auf Grundlage der beschränkten Nutzlast des eingesetzten UA und die dadurch reduzierte Auswahl an zur Verfügung stehenden Sensoren wur-

de ein Konzept der Sensordatenfusion dargelegt, welches eine kooperative und darauf aufsetzend konkurrierende Fusion von monogonen LiDAR- und monokularen Kameradaten nutzt, um eine dreidimensionale Sicht auf die Umwelt bereitzustellen, wobei das resultierende Sichtfeld maximal dem der Kamera entspricht und in Teilen über das des LiDAR hinausgeht.

Das monogone LiDAR und die monokulare Kamera generieren jeweils ausschließlich eine zweidimensionale Sicht auf die Umwelt. Das LiDAR generiert eine Schnittebene durch die betrachteten Objekte. Form und Ausdehnung der Objekte ober- und unterhalb dieser Ebene sind unbekannt. Die Bilder der Kamera sind skalierungsinvariant, so dass ohne absolutes Vorwissen nur relative Aussagen über Objekte im Bild erhoben werden können. Eine alleinige Verwendung eines der Sensoren zur Hindernisdetektion kann somit nicht empfohlen werden.

Die kooperative Fusion kombiniert absolute LiDAR- mit relativen Kameradaten von zuvor identifizierten korrespondierende Features, um die dreidimensionale Lage und Form von Objekten zu approximieren, welche ober- bzw. unterhalb der LiDAR-Ebene liegen. Bei den ermittelten Approximationen handelt es sich um planare Vereinfachungen der in der Kamerasicht betrachteten Objektoberflächen.

Eine planare Repräsentation einer Oberfläche spiegelt nur einen geringen Anteil der Objektoberflächen in natürlichen Umgebungen wider. Um eine realistische Approximation der betrachteten Oberflächen zu erhalten, wird eine konkurrierende Fusion der Approximationen mittels eines probabilistischen Belegtheitsgitters vorgestellt, wobei die planare Repräsentation über die Zeitschritte in das Belegtheitsgitter überführt werden, um über die Betrachtung der Objekte aus unterschiedlichen Blickwinkeln die „wahre“ Objektoberfläche heraus zu modellieren. Die probabilistische Kombination nutzt für die Inklusion der planaren Approximationen ein inverses Sensormodell, welches die Annahmen der kooperativen Fusion berücksichtigt, um das Herausmodellieren zu fokussieren.

Schlussendlich liegt eine metrische, diskretisierte Hindernisrepräsentation vor, welche den Raum, der durch die detektierten Hindernisse belegt ist, je nach Auflösung der Repräsentation realistisch wiedergibt und an eine Umsetzung einer Hindernisvermeidung als Weltmodell übergeben werden kann.

Die Funktionstüchtigkeit der vorgestellten Konzepte und deren Umsetzungen konnte anhand simulativ und real generierter Testdaten aufgezeigt werden. Die realen Kamera-, LiDAR- und GNSS-Daten wurden unter anderem mit einem unbemannten Hubschrauberversuchsträger generiert. Die Validierung mittels eines realen Testszenarios zeigte, dass das

vorgestellte Konzept eine umfassendere 3D-Sicht auf die Umwelt ermöglicht und dass für diesen Datensatz eine Steigerung der detektierten Hindernisse um 140% gegenüber der alleinigen Nutzung eines Laserscanners erreicht werden konnte.

Darüber hinaus erfolgte eine erfolgreiche Detailanalyse der dargelegten extrinsischen Kalibrierung und der bedingten Integration von Messungen in das Belgtheitsgitter, wobei auch in diesen Punkten Steigerungen der Genauigkeit bzw. Verhinderung der ungewollten Freizeichnung aufgrund der Diskretisierung betrachtet werden konnten.

In der vorliegenden Arbeit wurde ein statischer Aufbau von Kamera und LiDAR betrachtet. Mit diesem können Überhänge mit zurückgesetzten Rückwänden, wie Dachüberhänge, nicht detektiert werden, wenn die LiDAR die zurückgesetzte Wand beleuchten. Um dieses Problem zu beheben, müssen weitere dreidimensionale Daten bereitgestellt werden.

Ohne eine Steigerung an der benötigten Nutzlast zu generieren, wäre etwa die Nutzung eines bewegungsbasierten Stereoansatzes umsetzbar, der auf die gesamte Kameradatenbasis angewendet wird. Dies würde eine weitere Quelle für Tiefendaten auch außerhalb der LiDAR-Sicht bereitstellen. Es sind aber bekannte Schwächen zu beachten, speziell bei einem Flug in Richtung der optischen Achse bei dem sehr geringe Änderungen der benötigten Feature entlang dieser zu erwarten sind.

Rotierende oder schwenkende monogone Lasersysteme, wie sie in anderen Arbeiten präsentiert wurden, wären ebenfalls in der Lage ein umfassenderes Sichtfeld bereitzustellen. Deren Abtastraten sind je nach Rotationsgeschwindigkeit allerdings eher gering. Eine Kombination der einzelnen Laserschnittebenen mit einem Kamerabild würde nach dem hier präsentierten Ansatz eine Steigerung der Aktualisierungsrate für diese Sichtpyramide der Kamera bedeuten.

Die dargelegte Arbeit sowie die vorgestellten Ansätze bieten somit Anknüpfungspunkte für nachfolgende Forschungen.

Literatur

- [1] M. D. Adams. *Sensor modelling, design and data processing for autonomous navigation*. Bd. 13. World Scientific series in robotics and intelligent systems. Singapore: World Scientific, 1999.
- [2] F.-M Adolf u. a. „An Unmanned Helicopter for Autonomous Flights in Urban Terrain“. In: *German Workshop on Robotics* (2009).
- [3] T. Alexe. „Ein unbekanntes Flugobjekt: Eigentlich soll die Polizeidrohne den Sicherheitskräften bei ihren Einsätzen helfen. 2012 flog sie bislang nur ganze fünf Einsätze. Klarer Fall für die Mottenkiste, finden die Grünen.“ In: *Sächsische Zeitung* 2012 (1.07.2012).
- [4] H. Aliakbarpour u. a. „An Efficient Algorithm for Extrinsic Calibration between a 3D Laser Range Finder and a Stereo Camera for Surveillance“. In: *Advanced Robotics 2009 ICAR 2009 ICAR 2009. International Conference on*, S. 1–6.
- [5] J. Amanatides und A. Woo. „A Fast Voxel Traversal Algorithm for Ray Tracing“. In: *In Eurographics '87*. 1987, S. 3–10.
- [6] F. Andert und L. Goormann. „Combined grid and feature-based occupancy map building in large outdoor environments“. In: *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. 2007, S. 2065–2070. ISBN: 978-1-4244-0912-9.
- [7] F. Andert. *Bildbasierte Umgebungserkennung für autonomes Fliegen: Zugl.: Braunschweig, Techn. Univ., Diss., 2011*. Als Ms. gedr. Bd. 2011,9. Forschungsbericht / Deutsches Zentrum für Luft- und Raumfahrt. Köln: DLR, 2011.
- [8] E. Angel. *Interactive computer graphics: A top-down approach using OpenGL* ®. 4th ed., internat. ed., [Nachdr.] Boston und Mass: Pearson/ Addison-Wesley, 2006. ISBN: 032131252X.
- [9] A. Armbruster. „Wenn die Drohne zweimal klingelt: Amazon kündigt an, in wenigen Jahren Pakete mit Drohnen auszuliefern zu wollen. Unrealistisch ist das nicht, Vorreiter gibt es längst. Aber wollen wir das auch?“ In: *Frankfurter Allgemeine Zeitung* 2013 (2.12.2013).
- [10] BBC News Scotland Business, Hrsg. *Aerial 'drone' firm Cyberhawk draws investors*. 24.06.2013.

- [11] H. Baltzakis, A. Argyros und P. Trahanias. „Fusion of range and visual data for the extraction of scene structure information“. In: *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. Bd. 4. 2002, S. 7–11. ISBN: 1051-4651.
- [12] D. Bellot, A. Boyer und F. Charpillat. „A new definition of qualified gain in a data fusion process: application to telemedicine“. In: *Information Fusion, 2002. Proceedings of the Fifth International Conference on*. Bd. 2. 2002, S. 865–872. ISBN: 0-9721844-1-4.
- [13] D. Bi und X. Lu. „A New Flexible Approach for Single Laser stripe profiler Calibration“. In: *IEEE Proceedings on Information and Automation (2008)*, S. 76–80.
- [14] N. Blum. *Algorithmen und Datenstrukturen*. Oldenbourg Wissenschaftsverlag, 2012. ISBN: 9783486714036.
- [15] W. Boehler, B. M. Vicent und A. Marbs. „Investigating Laser Scanner Accuracy“. In: *Proceedings of XIXth CIPA International Symposium*. 2003.
- [16] J. Borenstein und Y. Koren. „The vector field histogram-fast obstacle avoidance for mobile robots“. In: *Robotics and Automation, IEEE Transactions on* 7.3 (1991), S. 278–288.
- [17] J. E. Bresenham. „Algorithm for Computer Control of a Digital Plotter“. In: *IBM Syst. J.* 4.1 (1965), S. 25–30.
- [18] R. R. Brooks und S. S. Iyengar. *Multi-sensor fusion: Fundamentals and applications with software*. Upper Saddle River N.J.: Prentice Hall PTR, 1998. ISBN: 0139016538.
- [19] D. C. Brown. „Close-range camera calibration“. In: *PHOTOGRAMMETRIC ENGINEERING* 37.8 (1971), S. 855–866.
- [20] J. Canny. „A Computational Approach to Edge Detection“. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI*-8.6 (1986), S. 679–698.
- [21] J. S. Dittrich, A. Bernatz und F. Thielecke. „Intelligent systems research using a small autonomous rotorcraft testbed“. In: *2nd AIAA "Unmanned Unlimited" Systems, Technologies and Operations —. Aerospace* (15 - 18 September 2003), S. 1–7.
- [22] S. F. El-Hakim. „Real-time image metrology with CCD cameras“. In: *Photogrammetric engineering and remote sensing* 52.11 (1986), S. 1757–1766+.
- [23] A. Elfes. „Sonar-based real-world mapping and navigation“. In: *Robotics and Automation, IEEE Journal of* 3.3 (1987), S. 249–265.

- [24] A. Elfes. „Using occupancy grids for mobile robot perception and navigation“. In: *Computer* 22.6 (1989), S. 46–57.
- [25] D. Elsner. „Volumetric Reconstruction of Objects and Scenes Using Range Images“. In: *Digital signal processing* (1999), Vol. 9, No. 2 (1999), 120–135.
- [26] J. H. G. Ender. *Hochauflösende bildgebende Radarverfahren mit elektronisch gesteuerten Gruppenantennen*. Paderborn: Ferdinand Schöningh Paderborn, 2003.
- [27] E. K. Forkuo und B. King. „Automatic fusion of photogrammetric imagery and laser scanner point clouds“. In: *In International Archives of Photogrammetry and Remote Sensing, vol. XXXV, part B4*. 2005, S. 921–926.
- [28] D. Fox, Wolfram Burgard und S. Thrun. „Markov Localization for Mobile Robots in Dynamic Environments“. In: *Journal of Artificial Intelligence Research* 11 (1999), S. 391–427.
- [29] D. Frost. *Implementierung und Vergleich von Methoden zur Kamerakalibrierung für ein Stereomeßsystem: Unveröffentl. Diplomarbeit*. Fachhochschule Oldenburg/Ostfriesland/Wilhelmshaven, 2008.
- [30] J. Galilea u. a. „Calibration of a High-Accuracy 3-D Coordinate Measurement Sensor Based on Laser Beam and CMOS Camera“. In: *Instrumentation and Measurement, IEEE Transactions on* 58.9 (2009), S. 3341–3346.
- [31] M. C. Garcia-Alegre u. a. „Real-Time Fusion of Visual Images and Laser Data Images for Safe Navigation in Outdoor Environments“. In: *Sensor Fusion - Foundation and Applications*. Hrsg. von C. Thomas. InTech, 2011, S. 121–138. ISBN: 978-953-307-446-7.
- [32] H. Goldstein und G. Gliemann. *Klassische Mechanik*. Aula-Verlag, 1991. ISBN: 9783891045145.
- [33] R. C. Gonzalez. *Digital image processing*. 3. Aufl. Upper Saddle River und NJ: Pearson/Prentice Hall, 2008. ISBN: 978-0-13-168728-8.
- [34] D. M. Green und J. A. Swets. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- [35] S. Grzonka, G. Grisetti und W. Burgard. „Towards a navigation system for autonomous indoor flying“. In: *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on* (12-17 May 2009), S. 2878–2883.

- [36] S. Grzonka, G. Grisetti und W. Burgard. „A Fully Autonomous Indoor Quadrotor“. In: *Robotics, IEEE Transactions on* 28.1 (2012), S. 90–100.
- [37] HOKUYO AUTOMATIC. *Scanning Laser Range Finder UTM-30LX/LN Specification*. Hrsg. von HOKUYO AUTOMATIC. Japan.
- [38] D. L. Hall und J. Llinas. *Handbook of multisensor data fusion*. The electrical engineering and applied signal processing series. Boca Raton: CRC Press, 2001. ISBN: 0849323797.
- [39] J. Hancock. „Laser Intensity-Based Obstacle Detection and Tracking“. Diss. Pittsburgh und Pennsylvania: Carnegie Mellon University, 26.01.1999.
- [40] M. Hebert und E. Krotkov. „3-D measurements from imaging laser radars: how good are they?“ In: *Intelligent Robots and Systems '91. 'Intelligence for Mechanical Systems, Proceedings IROS '91. IEEE/RSJ International Workshop on*. 1991, S. 359–364. ISBN: 0-7803-0067-X.
- [41] T. Hermes. *Digitale Bildverarbeitung: Eine praktische Einführung*. München [u.a.]: Hanser, 2005. ISBN: 3-446-22969-8.
- [42] R. Hohmann. „Fusion von Laserscanner- mit Stereokameradaten in einem gemeinsamen Weltmodell“. Masterthesis. Schmalkalden: Fachhochschule Schmalkalden, 19.09.2012.
- [43] A. Hornung u. a. „OctoMap: an efficient probabilistic 3D mapping framework based on octrees“. In: *Autonomous robots* 34.3 (2013), S. 189–206.
- [44] S. Hrabar. „3D path planning and stereo-based obstacle avoidance for rotorcraft UAVs“. In: *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on* (22-26 Sept. 2008), S. 807–814.
- [45] S. Hrabar u. a. „Combined optic-flow and stereo-based navigation of urban canyons for a UAV“. In: *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*. 2005, S. 3309–3316.
- [46] S. Hrabar. „An evaluation of stereo and laser-based range sensing for rotorcraft unmanned aerial vehicle obstacle avoidance“. In: *Journal of Field Robotics* 29.2 (2012), S. 215–239.
- [47] S. Hrabar und G. Sukhatme. „Vision-based navigation through urban canyons“. In: *Journal of Field Robotics* 26.5 (2009), S. 431–452.

- [48] IBERDROLA, Hrsg. *Iberdrola and CDTI invest 500,000euro in Salamanca-based company ARBÓREA INTELLBIRD, S.L.* 2013. URL: <http://www.suasnews.com/2013/06/23318/iberdrola-to-use-uavs-for-power-lines-monitoring/>.
- [49] B. Jähne. *Digitale Bildverarbeitung: Mit 155 Übungsaufgaben.* 6., überarb. und erw. Aufl. Berlin: Springer, 2005. ISBN: 3540249990.
- [50] F. Jorhden. „Fusion von Laserscanner- und Monokameradaten für eine präsumtive 3D-Sicht“. Bachelorthesis. Magdeburg: Otto von Guericke Universität Magdeburg, 15.02.2013.
- [51] I. T. Jolliffe. *Principal component analysis.* 2. ed. Springer series in statistics. New York und NY: Springer, 2002. ISBN: 0-387-95442-2.
- [52] L. A. Klein. *Sensor and data fusion concepts and applications.* 2. ed. Bd. 35. Tutorial texts in optical engineering. Bellingham und Wash.: SPIE, 1999. ISBN: 0819432318.
- [53] G. K. Kraetzschmar, G. P. Gassull und K. Uhl. „Probabilistic Quadtrees for Variable-Resolution Mapping of Large Environments“. In: *Eds.), Proceedings of the 5th IFAC/EURON.* 2004.
- [54] S. Krause und R. Evert. „Remission based improvement of extrinsic parameter calibration of camera and laser scanner“. In: *Control Automation Robotics Vision (ICARCV), 2012 12th International Conference on.* 2012, S. 829–834.
- [55] S. Krause. „Multi-Purpose Environment Awareness Approach for Single Line Laser Scanner in a Small Rotorcraft UA“. In: *Journal of Intelligent & Robotic Systems* 65 (2012), S. 587–601.
- [56] J. Lachman. „Im Drohnenfieber: Die fliegenden Roboter eröffnen einen Milliardenmarkt, auf dem sich auch deutsche Anbieter etablieren wollen.“ In: *Die Zeit* 2014.N^circ 2 (13.01.2014).
- [57] R. Lenz. „Linsenfehlerkorrigierte Eichung von Halbleiterkameras mit Standarobjektiven für hochgenaue 3D-Messungen in Echtzeit“. In: *Mustererkennung 1987, 9. DAGM-Symposium.* Springer-Verlag, 1987, S. 212–216. ISBN: 3-540-18375-2.
- [58] K. Levenberg. „A method for the solution of certain problems in least squares“. In: *Quart. Applied Math.* 2 (1944), S. 164–168.
- [59] G. Li u. a. „An algorithm for extrinsic parameters calibration of a camera and a laser range finder using line features“. In: *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on.* 2007, S. 3854–3859. ISBN: 978-1-4244-0912-9.

- [60] S. Lorenz. „Adaptive nichtlineare Regelung zur automatischen Flugbereichserweiterung für den Technologiedemonstrator ARTIS“. Diss. Technische Universität Carolo-Wilhelmina: Braunschweig, 2010.
- [61] T. Luhmann. „Erweiterte Verfahren zur geometrischen Kamerakalibrierung in der Nahbereichsphotogrammetrie“. Diss. München: Technische Universität Dresden, 2010.
- [62] D. W. Marquardt. „An algorithm for least-squares estimation of non-linear parameters“. In: *SIAM Journal on Applied Mathematics* 11.2 (1963), S. 431–441.
- [63] D. Meagher. *Geometric modeling using octree encoding*. Bd. IPL-TR-81-005. Reports / RENSELAER POLYTECHNIC INST TROY NY IMAGE PROCESSING LAB.
- [64] C. Mei und P. Rives. „Calibration between a Central Catadioptric Camera and a Laser Range Finder for Robotic Applications“. In: *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, title=Calibration between a central catadioptric camera and a laser range finder for robotic applications. 2006, S. 532–537.
- [65] T. Merz und F. Kendoul. „Beyond visual range obstacle avoidance and infrastructure inspection by an autonomous helicopter“. In: *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. 2011, S. 4953–4960. ISBN: 978-1-61284-454-1.
- [66] P. Moghadam, W. S. Wijesoma und J. F. Dong. „Improving path planning and mapping based on stereo vision and lidar“. In: *Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on*. 2008, S. 384–389. ISBN: 978-1-4244-2286-9.
- [67] H. P. Moravec. „Sensor Fusion in Certainty Grids for Mobile Robots“. In: *Sensor Devices and Systems for Robotics*. Hrsg. von A. Casals. Bd. 52. NATO ASI Series. Springer Berlin Heidelberg, 1989, S. 253–276. ISBN: 978-3-642-74569-0.
- [68] H. P. Moravec und A. Elfes. „High resolution maps from wide angle sonar“. In: *Robotics and Automation. Proceedings. 1985 IEEE International Conference on*. Bd. 2. 1985, S. 116–121.
- [69] D. Murray und C. Jennings. „Stereo vision based mapping and navigation for mobile robots“. In: *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*. Bd. 2. 1997, S. 1694–1699. ISBN: 0-7803-3612-7.

- [70] O. Naroditsky, A. Patterson und K. Daniilidis. In: *Robotics and Automation (ICRA), 2011 IEEE International Conference on, title=Automatic alignment of a camera with a line scan LIDAR system*. 2011, S. 3429–3434.
- [71] Y. Okubo, C. Ye und J. Borenstein. „Characterization of the Hokuyo URG-04LX laser rangefinder for mobile robot obstacle negotiation“. In: *Proc. SPIE 7332* (2009).
- [72] P. Payeur u. a., Hrsg. *Probabilistic octree modeling of a 3D dynamic environment*. Bd. 2. 1997.
- [73] T. Peynot und A. Kassir. „Laser-camera data discrepancies and reliable perception in outdoor robotics“. In: *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. 2010, S. 2625–2632. ISBN: 978-1-4244-6674-0.
- [74] U. Raschke und J. Borenstein. „A comparison of grid-type map-building techniques by index of performance“. In: *Robotics and Automation, 1990. Proceedings., 1990 IEEE International Conference on*. 1990, S. 1828–1832. ISBN: 0-8186-9061-5.
- [75] Y. Reshetyuk. „Self-calibration and direct georeferencing in terrestrial laser scanning“. Diss. KTH, Geodesy, 2009.
- [76] A. Rosenfeld und de la Torre, P. „Histogram concavity analysis as an aid in threshold selection“. In: *Systems, Man and Cybernetics, IEEE Transactions on SMC-13.2* (1983), S. 231–235.
- [77] J. Santolaria u. a. „A one-step intrinsic and extrinsic calibration method for laser line scanner operation in coordinate measuring machines“. In: *Measurement Science and Technology 20* (2009).
- [78] D. Scaramuzza, A. Harati und R. Siegwart. „Extrinsic self calibration of a camera and a 3D laser range finder from natural scenes“. In: *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. 2007, S. 4164–4169. ISBN: 978-1-4244-0912-9.
- [79] S. Scherer. „Low-Altitude Operation of Unmanned Rotorcraft“. Diss. Pittsburgh und PA: The Robotics Institute, Carnegie Mellon University, 2011.
- [80] S. Scherer u. a. „Flying Fast and Low Among Obstacles: Methodology and Experiments“. In: *Int. J. Rob. Res 27* (2008), S. 549–574.
- [81] M. Scholz. „Optimierung einer 3D-Umgebungsmodellierung mittels Octree am Beispiel von Laserdaten“. Diplomarbeit. Münster: Westfälische Wilhelms-Universität, 2012.

- [82] Sick AG. *Data sheet LMS 200: Betriebsanleitung, Technische Beschreibung, Telegrammliste*. Hamburg, 2005.
- [83] Sick AG, Hrsg. *SICK LD-MRS data sheet: Laser Measurement Technology LD-MRS / Outdoor / Long-Range LD-MRS 400001*. Waldkirch, 2010.
- [84] Sick AG, Hrsg. *Sick LMS291 data sheet: Laser Measurement Technology LMS2xx / LMS291 / Outdoor / Mid-Range LMS291-S05*. Waldkirch, 2010.
- [85] C. Stachniss. „Exploration and Mapping with Mobile Robots“. Diss. Freiburg im Breisgau: Albert-Ludwigs-Universität, 26.04.2006.
- [86] Teck Chew Ng, J. Guzman und Jin Chang Tan. „Development of a 3D LADAR system for autonomous vehicle guidance“. In: *Robotics, Automation and Mechatronics 2004 IEEE Conference 1 (2004)*, S. 13–18.
- [87] S. Thrun. „Learning occupancy grids with forward models“. In: *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*. Bd. 3. 2001, S. 1676–1681. ISBN: 0-7803-6612-3.
- [88] S. Thrun, W. Burgard und D. Fox. *Probabilistic Robotics*. Intelligent robotics and autonomous agents series. Mit Press, 2005. ISBN: 978-02-6220-162-9.
- [89] S. Thrun und A. Bücken. „Integrating Grid-based and Topological Maps for Mobile Robot Navigation“. In: *Proceedings of the Thirteenth National Conference on Artificial Intelligence - Volume 2. AAAI'96*. AAAI Press, 1996, S. 944–950. ISBN: 0-262-51091-X.
- [90] S. Thrun, M. Diel und D. Hähnel. „Scan alignment and 3d surface modeling with a helicopter platform“. In: *In Proceedings of the International Conference on Field and Service Robotics*. 2003.
- [91] C. Tomasi und R. Manduchi. „Bilateral filtering for gray and color images“. In: *Computer Vision, 1998. Sixth International Conference on*. 1998, S. 839–846. ISBN: 81-7319-221-9.
- [92] R. Y. Tsai. „An efficient and accurate camera calibration technique for 3D machine vision“. In: *Computer Vision and Pattern Recognition*. 1986.
- [93] Velodyne Lidar Inc. *High Definition Lidar HDL-64E S2*. 2008.
- [94] K. Vössing. „Drohnen werden sicherer sein als Autos: TU-Studenten haben einen Flugroboter entwickelt. 2014 soll er einmalig Pizza transportieren und irgendwann Medikamente.“ In: *Braunschweiger Zeitung* 2013. Braunschweig (13.12.2013).

- [95] L. Weibking. *Entwicklung eines zentimetergenauen mehrdimensionalen Nahbereichs-Navigationssystems*. Bd. 994. Fortschr.-Ber. VDI Reihe 8. Düsseldorf: VDI Verlag, 2003.
- [96] M. Whalley u. a. *Design and Flight Test Results for a Hemispherical LADAR Developed to Support Unmanned Rotorcraft Urban Operations Research*. 28.05.2008.
- [97] O. Wulf. *Virtuelle 2D-Scans - Ein Verfahren zur echtzeitfähigen Roboternavigation mit dreidimensionalen Umgebungsdaten*. Berlin: dissertation.de, 2008.
- [98] K. M. Wurm u. a. „OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems“. In: *Proc. of the ICRA 2010 workshop on best practice in 3D perception and modeling for mobile manipulation* (2010).
- [99] Q. Zhang und R. Pless. „Extrinsic calibration of a camera and laser range finder (improves camera calibration)“. In: *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*. Bd. 3. 2004, S. 2301–2306.
- [100] Z. Zhang. „Flexible camera calibration by viewing a plane from unknown orientations“. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. Bd. 1. 1999, S. 666–673.