

Geodätisch-geophysikalische Arbeiten in der Schweiz

(Fortsetzung der Publikationsreihe
«Astronomisch-geodätische Arbeiten in der Schweiz»)

herausgegeben von der

Schweizerischen Geodätischen Kommission
(Organ der Akademie der Naturwissenschaften Schweiz)

**Fünfundachtzigster Band
Volume 85**

**Echtzeit-Georegistrierung von
Videodaten mit Hilfe von Navigations-
sensoren geringer Qualität und
digitalen 3D-Landschaftsmodellen**

Hannes Eugster

2012

Adresse der Schweizerischen Geodätischen Kommission:

Institut für Geodäsie und Photogrammetrie
Eidg. Technische Hochschule Zürich
ETH Zürich
8093 Zürich
Switzerland

Internet: <http://www.sgc.ethz.ch>

ISBN 978-3-908440-29-1

Redaktion des 85. Bandes:
Dr. H. Eugster, J. Müller-Gantenbein, A. Geiger
Druck: Print-Atelier ADAG, Zürich

VORWORT

Der kombinierte Einsatz von Globalen Satellitennavigationssystemen (GNSS) und von Trägheitsnavigationssystemen (INS) zur direkten Georeferenzierung von Geosensoren, wie beispielsweise digitalen Luftbildkameras oder luftgestützten Laserscannern, ist heute Stand der Technik und wird in der Praxis routinemässig eingesetzt. Um Georeferenzierungsgenauigkeiten im Sub-Dezimeterbereich zu erreichen, kommen bisher ausschliesslich hochwertige GNSS/INS-Messsysteme mit den entsprechenden Kosten- und Gewichtsfolgen zum Einsatz.

Markante Fortschritte in der Mikrosystem-, Regelungs- und Kommunikationstechnik haben zu leichten und kostengünstigen Sensoren und Plattformen geführt. So weisen beispielsweise heutige Smartphones eine Vielzahl von Sensoren wie etwa MEMS-basierte Beschleunigungs- und Drehratensensoren sowie hochaufgelöste Miniaturkameras auf. Dies begünstigt wiederum die Entwicklung leichter unbemannter Flugsysteme (Mini und Micro UAVs) und deren rasch zunehmende Nutzung für Kartierungs- und Geomonitoring-Anwendungen – bis jetzt allerdings mit einer um eine Grössenordnung ungenaueren direkten Georeferenzierung.

Die Arbeit von Hannes Eugster überwindet diesen Nachteil durch einen neuartigen integrierten Georeferenzierungsansatz. Er hat dazu Ansätze aus der Computer Vision, der Photogrammetrie und der geodätischen Statistik aufgegriffen und diese auf sehr innovative Art und Weise weiter entwickelt. Kernidee der Arbeit ist die Nutzung existierender 3D-Geometriedaten, insbesondere von 3D-Stadtmodellen, als Basis für eine integrierte Georeferenzierung von mobil erfassten Videodaten. Die massgebenden Beiträge dieser Arbeit können wie folgt umrissen werden: die Entwicklung eines relationalen Bild-zu-3D-Stadtmodell-Zuordnungsverfahrens in Kombination mit einer neuen kontinuierlichen, robusten Bündelblock-Triangulation sowie deren Integration in ein komplementäres Kalman-Filter-Design zur kontinuierlichen Georeferenzierung und Fehlerzustandsschätzung. Ein weiteres innovatives Element ist die Art des Zugriffs auf die 3D-Referenzgeometrien als Basis für die integrierte Georegistrierung. Hannes Eugsters Ansatz baut erstmals auf einem Virtuellen Globus auf, der hier nicht nur als Visualisierungswerkzeug dient, sondern als streambare 3D-Geodatenbasis bzw. semantische 3D Geospatial Query Engine.

Mit der Implementierung eines voll funktionsfähigen Prototypsystems zur integrierten (Nah-) Echtzeitgeoregistrierung von mobil erfassten Videodaten hat Hannes Eugster das Funktionieren seines integrierten Verfahrens eindrücklich demonstriert und bewiesen. Durch die Nutzung existierender 3D-Referenzgeometrien erlaubt es der neue Lösungsansatz, auch mit kostengünstigen Sensoren und leichten Sensorplattformen wie Micro UAVs eine vollautomatische Georeferenzierung in (Nah-)Echtzeit in der Grössenordnung der Genauigkeit der verfügbaren 3D-Referenzgeometrie zu erzielen. Damit eröffnet die Arbeit von Hannes Eugster auch viele neue Anwendungsmöglichkeiten, beispielsweise für zukünftige qualitativ hochstehende mobile Augmented Reality Anwendungen im städtischen Raum mit Helmkameras oder Smartphones.

Die Schweizerische Geodätische Kommission (SGK) dankt dem Verfasser, Herrn Dr. Eugster, für den wertvollen Beitrag zur Weiterentwicklung neuer geodätischer Verfahren. Der SCNAT danken wir für die Übernahme der Druckkosten.

Prof. Dr. Stephan Nebiker
Institut Vermessung und Geoinformation
FHNW Fachhochschule Nordwestschweiz

Prof. Dr. A. Geiger
ETH Zürich
Präsident der SGK

PREFACE

La mise-en-œuvre combinée de la navigation par satellites (GNSS) et par inertie (INS) pour la référence géographique directe de senseurs comme, par exemple, les caméras aériennes digitales ou les « scanner » laser embarqués, représente l'état actuel de la technique et qui sera appliquée d'une façon routinière dans la pratique. Afin d'atteindre une précision sub-décimétrique dans la géo-référenciation sont mis en œuvre exclusivement des systèmes de mesures chers avec des coûts et des poids en découlant.

De marquants progrès dans les techniques des microsystemes, de réglage et de communication ont conduits à la fabrication de senseurs et de plateformes meilleur marché et plus légers. Ainsi, par exemple, les « Smartphones » actuels contiennent un grand nombre de senseurs comme des accéléromètres ou des détecteurs de rotation, basés sur des MEMS, et des caméras miniatures à hautes résolutions. D'autre part apparait le développement de systèmes aériens sans équipage (mini et micro UAV) et la rapide croissance de l'utilisation de tels systèmes pour la cartographie et pour des applications de géo-surveillance remplace ce qui jusqu'alors était fait par une géo-référenciation directe d'une grande imprécision.

Le travail de Monsieur Hannes Eugster surmonte ces inconvénients à l'aide d'une nouvelle mise-en-œuvre de la géo-référenciation. Il a, de plus, d'une manière très innovatrice, saisi et développé des parties venant de la vision par ordinateur, de la photogrammétrie et de la statistique géodésique. L'idée centrale de ce travail est l'utilisation de données tridimensionnelles existantes, particulièrement des modèles urbains 3-D, comme base pour une intégration géo-référencée de saisies de données de vidéo mobiles. Les contributions déterminantes de ce travail peuvent être décrites comme suit : Le développement d'une procédure relationnelle de classement d'images de modèles urbains tridimensionnels. Ceci en combinaison avec une nouvelle, robuste et continue d'estimation en bloc des paramètres de triangulation (Bündelblock-Triangulation), de même que son intégration dans la création de filtres de Kalman complémentaires pour la géo-référenciation et l'évaluation des erreurs en continu. Un autre élément innovatif est la méthode d'accès à la géométrie tridimensionnelle de référence comme base pour le géo-enregistrement intégré. L'approche de Monsieur Hannes Eugster construit pour la première fois sur un globe virtuel, ne servant pas seulement comme outil de visualisation, mais aussi comme base de données 3D et aussi comme un engin de recherche sémantique spatiale 3D en ligne (Geospatial Query Engine)

Avec l'implémentation d'un prototype de ce système complètement fonctionnel pour l'intégration en temps réel de géo-enregistrements de vidéo mobile, Monsieur Hannes Eugster a démontré et prouvé le fonctionnement de sa procédure intégrée. A l'aide de l'utilisation de géométries de références tridimensionnelles la nouvelle approche permet aussi une complète géo-referenciation en ligne avec de senseurs meilleurs marchés et des plateformes légères, comme par exemple des micros UAV, et d'atteindre les géométries 3D disponibles avec leurs ordres de grandeurs de précision. Avec ceci le travail de Monsieur Hannes Eugster ouvre aussi de nouvelles possibilités d'applications, par exemple pour de futures applications de haut niveau de réalité augmentée (Augmented Reality) dans l'espace urbain avec des caméras frontales ou des téléphones portables.

La Commission Géodésique Suisse (CGS) est reconnaissante envers l'Académie Suisse des Sciences Naturelles (ScNAT) pour avoir pris à sa charge les coûts d'impression du présent manuscrit.

Prof. Dr. Stephan Nebiker
Institut Vermessung und Geoinformation
FHNW Fachhochschule Nordwestschweiz

Prof. Dr. A. Geiger
ETH Zürich
Président de la CGS

FOREWORD

Direct georeferencing of geosensors like digital aerial imaging or airborne laser scanners using both, the global navigation satellite system (GNSS) and the inertial navigation system (INS), is today's state-of-the-art technology and routinely used. In order to reach a georeferencing accuracy below the decimeter level, high quality GNSS and INS measurement devices are used exclusively. Such equipment is inherently connected with a substantial financial expense and it is also relatively heavy.

Substantial progress in the development of microsystems-, control-, and communication techniques lead to smaller, lighter, and less expensive sensors and measurement platforms. For example, modern Smartphones contain a number of different sensors like MEMS based acceleration and rotation-rate sensors as well as high resolution miniature cameras. This also favors the development towards light unmanned flight systems (Mini and Micro UAVs) along with its increasing use in field mapping and geomonitoring applications. However, the accuracy of these miniature systems is about one order of magnitude lower than what can be obtained with today's state-of-the-art measurement systems.

As the work of Hannes Eugster shows, this drawback can be resolved using a new, integrated georeferencing module. In his work, principles of computer vision, photogrammetry and geodetic statistics were used and further developed in a very innovative way. His principle idea is to use existing 3D geometry data, especially 3D city models, as a basis for integrated georeferencing of mobile captured video data. The main contribution of this work can be summarized as follows: Development of a relational image-to-3D-city-model-assignment-procedure in combination with a new continuous and robust batch-block-triangulation as well as its integration into a complementary Kalman filter design used for a continuous georeferencing and error estimation. Another innovative aspect is the way of accessing the 3D reference geometries as a basis for the integrated georegistration. For the first time, a virtual globe is used not only as a visualising tool but as a streamable 3D geodata basis, or a semantic 3D geospatial query engine, respectively.

By implementing a fully operational prototype system for the integrated (near-) real-time georegistration of mobile acquired video data, Hannes Eugster impressively demonstrated the functionality of his integrated procedure. When using existing 3D reference geometries, the new procedure allows to automatically performing georeferencing in (near-) real-time with light and low cost sensor platforms like micro UAVs on an accuracy level comparable to the existing 3D reference geometries. The work of Hannes Eugster opens a series of new application possibilities, for example in the field of future high quality mobile augmented reality applications within cities and helmet cameras or Smartphones.

The Swiss Geodetic Commission (SGC) is grateful to the Swiss Academy of Sciences (ScNAT) for covering the printing costs of this volume.

Prof. Dr. Stephan Nebiker
Institut Vermessung und Geoinformation
FHNW Fachhochschule Nordwestschweiz

Prof. Dr. A. Geiger
ETH Zurich
President of SGC

Zusammenfassung

Kostengünstige und leichtgewichtige Navigationssensoren erlauben heute die kontinuierliche Positionierung und Ausrichtung mobiler Trägerplattformen zur Erfassung beliebiger Video- und Bilddaten. Dabei hängt die erreichbare direkte Georeferenzierungsgenauigkeit von der Qualität der eingesetzten Navigationssensorik ab. Für viele Anwendungen wie beispielsweise *Eco-Localisation* Aufgaben auf Basis von MEMS-INS und GNSS-Navigationsempfängern oder *Location Based Services* umgesetzt mit *Smartphones* ist die erreichbare Registrierungsgenauigkeit bezüglich bestehender Geodatenbeständen nicht ausreichend. Diese Arbeit präsentiert Konzepte und Ansätze welche die exakte Echtzeit-Georegistrierung mobil erfasster Video- und Bilddaten aufgrund näherungsweise bekannter Orientierungsdaten auf ein bestehendes digitales 3D-Stadtmodell ermöglichen. Die Umsetzung und Validierung der entwickelten Konzepte erfolgt konkret mittels Mini- und Mikrodrohnen erfasster Bild- und Videodaten. Diese Trägerplattformen ermöglichen insbesondere eine flexible und effiziente Bild- und Videodatenerfassung. Aus Gewichts- und Kostengründen kommen auf Mikro- und Minidrohnen für die Fluglagebestimmung lediglich INS- und GNSS-Sensoren geringer Qualität zum Einsatz, was für die Georeferenzierung der erfassten Geodaten für viele Anwendungen nicht genügt.

Das entworfene Lösungskonzept basiert auf einer zweistufigen Georeferenzierung. Die Grobregistrierung in der ersten Stufe erfolgt mit Hilfe der direkten Georeferenzierung basierend auf den zur Verfügung stehenden Navigationsdaten aus dem Fluglagekontroller. Ist die erzielbare Georegistrierungsgenauigkeit der direkten Lösung ungenügend, ermöglicht die anschliessende integrierte Georeferenzierung die Feinregistrierung. Dazu wird ein automatisches Bild-zu-Modell-Zuordnungsverfahren verwendet, welches die erfassten Bilddaten auf Basis der zur Verfügung stehenden Näherungslösung aus der ersten Registrierungsstufe in Beziehung zu den digitalen 3D-Stadtmodelldaten setzt. Basierend auf den im Bild zugeordneten Modellkanten werden wahlweise, mittels Einzelbildorientierung oder kontinuierlichem Bündelblock für einzelne Schlüsselbilder, unabhängig bestimmte äussere Orientierungsparameter der Videosequenz abgeleitet. Daraus schätzt ein komplementäres Kalman-Filter fortlaufend den verbleibenden systematischen Fehler der direkten Georeferenzierung. Die Kenntnis des aktuellen Fehlerzustands erlaubt anschliessend die Korrektur und Verbesserung der Georegistrierung. Das entwickelte Bild-zu-Modell-Zuordnungsverfahren ermöglicht insbesondere die Feinregistrierung von Bildaufnahmen auf bestehende 3D-Stadtmodelldaten ausgehend von einer näherungsweise bestimmten Aufnahmeposition und -ausrichtung. Weitergehend kann die vorgeschlagene integrierte Georeferenzierung als INS/GNSS/Vision-Datenfusionsansatz aufgefasst werden. Dieser erlaubt zum einen die Stützung der Navigationslösung mit Hilfe der zusätzlichen Bildinformation und

ermöglicht zum anderen die Bestimmung der Plattformlagedaten bezüglich bestehenden digitalen 3D-Modelldaten.

Der entwickelte Video-Georeferenzierungs-Prototyp dient der Validierung des vorgeschlagenen zweistufigen Georegistrierungsansatzes und den darin verwendeten Algorithmen. Parallel dazu wurde die erste Georegistrierungsstufe direkt in die operative Softwarelösung UAVision integriert, mit welcher die Echtzeitverarbeitung demonstriert werden kann. Die Funktionstüchtigkeit des vorgeschlagenen zweistufigen Georegistrierungsansatzes kann mittels Mini- und Mikrodrohnen aufgezeichneter realer Videosequenzen bestätigt werden. Durchgeführte Genauigkeitsuntersuchungen zeigen, dass der integrierte Ansatz eine Steigerung der Georegistrierungsgenauigkeit durchschnittlich um den Faktor drei im Vergleich zur direkten Georeferenzierungslösung ermöglicht. Abhängig von der Flugstabilität der verwendeten Aufnahmeplattform, der Qualität der Navigationssensorik sowie der verfügbaren Fluglagedaten-Updatefrequenz lassen sich mit dem vorgeschlagenen zweistufigen Ansatz Registrierungsgenauigkeiten für Minidrohnen im Bereich von 0.5m und für Mikrodrohnen von 1m im Objektraum erzielen.

Abstract

Today, low-cost and light-weight navigation sensors allow continuous positioning and attitude determination of mobile platforms, which can be used for geodata acquisition. The georeferencing accuracy which can be achieved depends directly on the quality of navigation sensors used. For many application scenarios, like eco-localisation problems realised with MEMS-INS in combination with GNSS navigation-grade receivers, or location-based services based on smartphones, the achievable registration accuracy with respect to existing geodatabase content is insufficient. This work develops and investigates concepts and approaches which allow real-time georegistration of mobile-captured video streams and single images based on coarse initial position and attitude information in combination with existing digital 3d city models. For the realisation and validation of the developed concepts and algorithms, video streams which have been captured with mini and micro unmanned aircraft are used. These platforms are particularly flexible and cost-effective for real-time geodata acquisition. Because of the limited payload capacity and costs of mini and micro unmanned aircraft, only low cost and quality navigation sensors are employed for flight data determination. However, these sensors also achieve a direct georeferencing accuracy which is insufficient for many mapping applications.

The proposed concept consists of a two-step georeferencing approach. The first step uses the direct georeferencing approach which is based on the available flight data from the flight controller being used for coarse registration. In applications where the achievable georegistration accuracy from the first step is insufficient, the following integrated georeferencing approach allows for accuracy improvement and precise image-registration. To this end, first a fully automatic image-to-model matching approach is used. This matching approach applies extracted edges from a single image to an existing digital 3d city model with the aid of the known coarse initial georeferencing solution. Based on the assigned 3d city model edges in each key frame of the image sequence, a resection based or a continuous bundle adjustment approach, allows the independent estimation of exterior orientations. The independently determined exterior orientation updates are subsequently used in a complementary Kalman filter which estimates the current error state of the direct georeferencing solution. Awareness of the current error state allows continuous correction and improvement of the georegistration. Especially the developed image-to-model matching algorithm enables the precise registration of images with respect to existing 3d city models based on coarse initial position and attitude information. Further, the proposed integrated georeferencing approach can be considered as an INS/GNSS/Vision data fusion approach, which supports on the one hand, a navigation solution with the aid of the additionally available imagery data. On the other hand, the proposed solution allows registration and navigation data estimation with respect to an existing geodetic reference system.

A video-georeferencing-prototype software was developed for validation of the proposed georegistration approach and testing of algorithms. At the same time, the first georeferencing step was implemented in the available UAVision software framework. With UAVision, real-time video georegistration can be demonstrated. Different video sequences, acquired with micro and mini unmanned aircraft and processed with the video-georeferencing-prototype software, show the serviceability of the proposed georegistration approach. Investigations show a georegistration accuracy improvement of a factor of three with the integrated georeferencing solution compared to the direct georeferencing solution. Depending on flight stability and flight platform, the quality of the integrated navigation sensors as well as the available flight state update frequency, a georegistration accuracy of 0.5m for mini and 1m for micro platforms are achievable with the proposed two-stage georegistration approach.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation.....	2
1.2	Ziele und Abgrenzung.....	5
1.3	Überblick über die Arbeit.....	6
2	Konzepte und Grundlagen	9
2.1	Konzepte zur Video-Georeferenzierung	10
2.1.1	Direkte und integrierte Georeferenzierung in der Photogrammetrie.....	10
2.1.2	Video Registration.....	11
2.1.3	Objektrekonstruktion aus Bildsequenzen.....	12
2.1.4	Simultaneous Localisation and Mapping.....	15
2.1.5	Zusammenstellung und Beurteilung.....	17
2.2	Grundlagen zur Bild-zu-3D-Stadtmodell-Zuordnung	19
2.2.1	Zuordnungsverfahren.....	19
2.2.2	Bild-zu-Modell-Zuordnung und Kanten-Matching.....	25
2.3	Grundlagen zur photogrammetrischen Einzel- und Mehrbildorientierung	30
2.3.1	Indirekte Georeferenzierung.....	31
2.3.2	Direkte Georeferenzierung.....	36
2.3.3	Integrierte Georeferenzierung.....	39
2.3.4	Gegenüberstellung der Bildorientierungsansätze.....	40
2.4	Grundlagen zur Kalman-Filter-basierten Sensordatenfusion	41
2.5	Basistechnologien	43
2.5.1	Virtuelle Globen.....	43
2.5.2	Unbemannte Flugsysteme.....	44
2.5.3	INS/GNSS-basierte Navigationssysteme.....	45
2.5.4	Videotechnik und -übertragung.....	47
3	Mobile Video-Georegistrierung	49
3.1	Architektur und Integrationskonzept	50
3.1.1	Ausgangslage und Voraussetzungen.....	50
3.1.2	Datenverarbeitungskette und Architektur.....	51
3.1.3	Zweistufige Video-Georegistrierung.....	53

3.2	Direkte Video-Georeferenzierung - Registrierungsstufe I	55
3.2.1	Umsetzung direkte Video-Georeferenzierung	55
3.2.2	Navigationsdaten geringer Qualität	62
3.3	Integrierte Video-Georeferenzierung - Registrierungsstufe II.....	67
3.3.1	Architektur des integrierten Lösungsansatzes	67
3.3.2	Bild-zu-Modell-Zuordnung.....	69
3.3.3	Bildorientierung und Online-Triangulation	85
3.3.4	Kontinuierliche Fehlerzustandsschätzung	92
4	Prototypanwendung und Umsetzung.....	99
4.1	Umsetzungskonzept und Programmsysteme.....	100
4.2	UAVision - Drohnenbasierte direkte Video-Georeferenzierung und -Integration	101
4.2.1	Entwicklungsumgebung.....	102
4.2.2	Video-Datenverarbeitung und -Integration	103
4.2.3	Funktionalität und Anwendungen	104
4.3	VGR-Prototyp - Drohnenbasierte integrierte Video-Georeferenzierung	105
4.3.1	Entwicklungsumgebung.....	106
4.3.2	Umsetzung Video-Georegistrierung	107
4.3.3	Steuerparameter und Prozessierungseinstellungen.....	109
5	Testflüge und Ergebnisse.....	111
5.1	Aufnahmeplattformen und Testgebiete	112
5.1.1	Mikro- und Minidrohnsysteme.....	112
5.1.2	Testgebiete und 3D-Stadtmodelle.....	114
5.2	Testsequenzen und Untersuchungen.....	116
5.2.1	Systemkalibrierung.....	116
5.2.2	Testsequenzen und Resultate	117
5.2.3	Untersuchungen	121
5.3	Beurteilung und Erfahrungen	128
5.3.1	Unzulänglichkeiten und Schwächen	128
5.3.2	Echtzeitfähigkeit	129
6	Fazit und Ausblick	131
6.1	Zusammenfassung	132
6.2	Empfehlungen	133
6.3	Anwendungsmöglichkeiten.....	134
6.4	Ausblick	135
Literatur		137

Anhang	149
A Parametrisierung Rotationsmatrizen	150
B Parameterschätzung nach der Methode der kleinsten Quadrate	152
B.1 Gauss-Newton Optimierung	154
B.2 Levenberg-Marquardt Optimierung	155
C Bildorientierung	157
C.1 Indirekte Georeferenzierung	157
C.2 Direkte Georeferenzierung.....	159
D Zeitdiskretes Kalman-Filter	160
E Chamfer-Algorithmus	161
F A* Algorithmus - informierte Baumsuche	163
G Kontinuierlicher Bündelblockausgleich	165
Dank	167

Konventionen und Abkürzungen

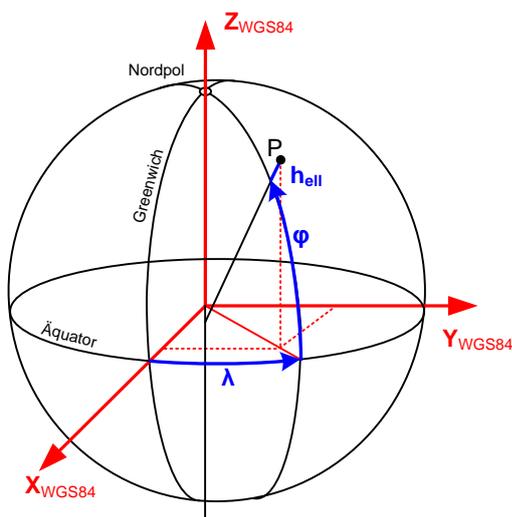
Konventionen

Fett gedruckte serifenlose Bezeichnungen wie z.B. \mathbf{EO}_{DG} stehen für Objekte oder Variablen in den Abbildungen. Eingeführte und verwendete englische Fachbegriffe werden in *kursiver* Schreibweise geschrieben. Diese Arbeit ist nach den Regeln der Schweizerischen Rechtschreibung verfasst.

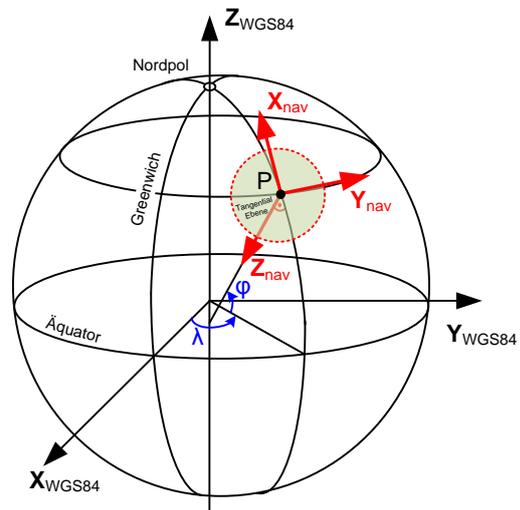
Koordinatensysteme

Zwei kartesische Koordinatensysteme können mit Hilfe einer Rotationsmatrix \mathbf{R} ineinander überführt werden. Dabei bezeichnet die Rotationsmatrix \mathbf{R}_a^b die Drehung vom Koordinatensystem a ins Koordinatensystem b .

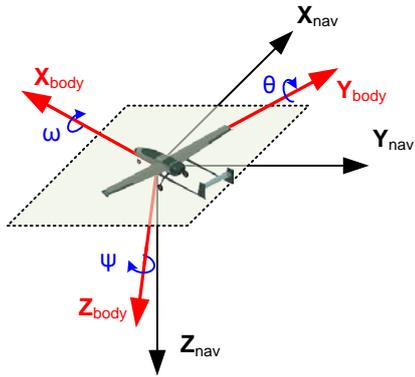
Geozentrisches kartesisches Koordinatensystem: (WGS 84)



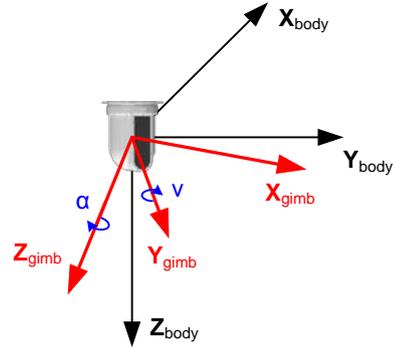
Topozentrisches Koordinatensystem: Navigationskoordinatensystem (nav)



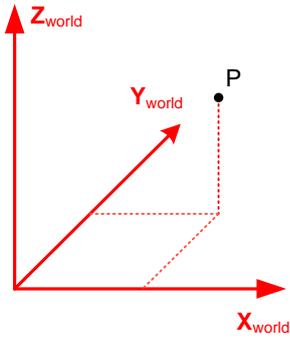
Trägerplattformsystem:
INS-Körperkoordinatensystem (*body*)



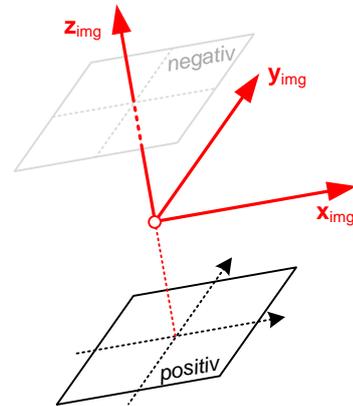
Sensoraufhängungssystem:
Gimbalsystem (*gimb*)



Übergeordnetes geodätisches Referenz-
koordinatensystem: z.B. Projektion einer
Landesvermessung (*world*)



Kamera- oder Sensorkoordinaten-
system: (*img* oder *sen*)



Abkürzungen

ABM	Flächenbasierte Zuordnung (engl. <i>Area Based Matching</i>)
AM	<i>Augmented Monitoring</i>
AR	<i>Augmented Reality</i>
AT	Aerotriangulation
BA	Bündelblockausgleichung (engl. <i>Bundle Adjustment</i>)
B-Rep	Randdarstellungsrepräsentation (engl. <i>Boundary Representation</i>)
CEP	<i>Circular Error Probable</i>
DG	Direkte Georeferenzierung
DGPS	differentielle Positionsbestimmung
DLT	Direkte Lineare Transformation
EFO	Energiefunktional-Optimierung (engl. <i>Energy Functional Optimisation</i>)
EKF	Erweitertes Kalman-Filter (engl. <i>Extended Kalman Filter</i>)
EO	Äussere Orientierung (engl. <i>Exterior Orientation</i>)
FBM	Merkmalsbasierte Zuordnung (engl. <i>Feature Based Matching</i>)
GHT	Generelle Hough-Transformation
GIS	Geografisches Informationssystem
GNSS	Globales Navigationssatellitensystem
GPS	US amerikanisches globales Satellitenpositionierungssystem
GUI	Graphische Benutzeroberfläche (engl. <i>Graphical User Interface</i>)
HT	Hough-Transformation
IMU	Inertiale Messeinheit (engl. <i>Inertial Measurement Unit</i>)
INS	Inertiales Navigationssystem
IO	Innere Orientierung (engl. <i>Interior Orientation</i>)
KAR	<i>Kinematic Ambiguity Resolution</i>
KF	Kalman-Filter
KLT	Kanade-Lucas-Tomasi <i>Feature Tracker</i>
LBS	<i>Location Based Services</i>

LiDAR	<i>Light Detection And Ranging</i>
LOD	Detailierungsstufe (engl. <i>Level of Detail</i>)
LSM	Kleinste Quadrate Anpassung (engl. <i>Least Squares Matching</i>)
LTC	engl. <i>Longitudinal Timecode</i>
LV95	Neue Landesvermessung der Schweiz 1995
MATLAB	<i>MATrix LABoratory</i>
MdkQ	Methode der kleinsten Quadrate (engl. <i>Least Squares Method</i>)
MEMS	<i>Micro-Electro-Mechanical Systems</i>
PAL	<i>Phase Alternating Line</i>
POI	<i>Point of Interest</i>
PPP	<i>Precise Point Positioning</i>
PR	Mustererkennung (engl. <i>Pattern Recognition</i>)
RANSAC	<i>RANdom SAmple Consensus</i>
RM	Relationale Zuordnung (engl. <i>Relational Matching</i>)
RTK	Relative GNSS-Positionierung (engl. <i>Real-time Kinematic</i>)
SEP	<i>Spherical Error Probable</i>
SfM	<i>Structure from Motion</i>
SLAM	<i>Simultaneous Localisation and Mapping</i>
SVD	Singulärwertzerlegung (engl. <i>Singular Value Decomposition</i>)
TLS	Terrestrisches Laserscanning
UA	<i>Unmanned Aircraft</i>
UAS	<i>Unmanned Aircraft System</i>
UKF	<i>Unscented Kalman-Filter</i>
UTC	Koordinierte Weltzeit (engl. <i>Universal Time Coordinated</i>)
VGR	Video-Georeferenzierungs-Prototyp
VITC	<i>Vertical Interval TimeCode</i>
WGS 84	<i>World Geodetic System 1984</i>
XML	<i>Extensible Markup Language</i>

1 Einleitung

1.1 Motivation

Leichtgewichtige und kostengünstige Navigationssensoren ermöglichen heute die Positions- und Ausrichtungsbestimmung unterschiedlicher Aufnahmesensorsysteme im Raum. Vielfach wird für die Bestimmung der Navigationsdaten MEMS (*Micro-Electro-Mechanical Systems*) basierte inertielle Messtechnik mit globaler Satellitennavigationstechnologie kombiniert. Damit lässt sich die Position und Ausrichtung mit einer Genauigkeit von wenigen Metern bzw. Grad kontinuierlich bestimmen. Der Einsatz dieser Navigationssensorik zusammen mit bildgebenden Sensoren und in Kombination mit bestehenden Geodatenbeständen ermöglicht die Umsetzung unterschiedlichster neuer Geoanwendungen wie beispielsweise bildbasierte *Location Based Services* (LBS) auf *Smartphones* oder bildgestützte *Ego-Localisation* in bebauten Regionen. Die erfolgreiche Umsetzung verlangt eine akkurate Registrierung der erfassten Bilddaten auf bestehende 3D-Landschaftsmodelldaten. Entsprechend werden Ansätze und Algorithmen benötigt, die eine Verbesserung der Georeferenzierungsgenauigkeit durch den zusätzlichen Einbezug der erfassten Bilddaten in Kombination mit den zur Verfügung stehenden Geobasisdatenbeständen gewährleisten.

Untersucht und gelöst werden soll diese Fragestellung anhand der folgenden konkreten Anwendung und Ausgangslage. Verschiedenste 3D-Geoinformationssysteme basierend auf virtuellen Globen (engl. *Virtual Globe*) stehen heute einem breiten Benutzendenkreis zur Verfügung. Trotz der Leistungsfähigkeit dieser Technologien in Bezug auf die effiziente Nutzbarmachung riesiger vorprozessierter Geodatenbestände weisen virtuelle Globen der aktuellen Generation auch Defizite auf. Meist sind die vorprozessierten Orthotexturdaten der digitalen Landschaftsmodelle nicht aktuell, was die Nutzung von 3D-Geoinformationssystemen für die Realisierung von Echtzeit- oder Entscheidungsunterstützungsanwendungen zurzeit erheblich einschränkt.

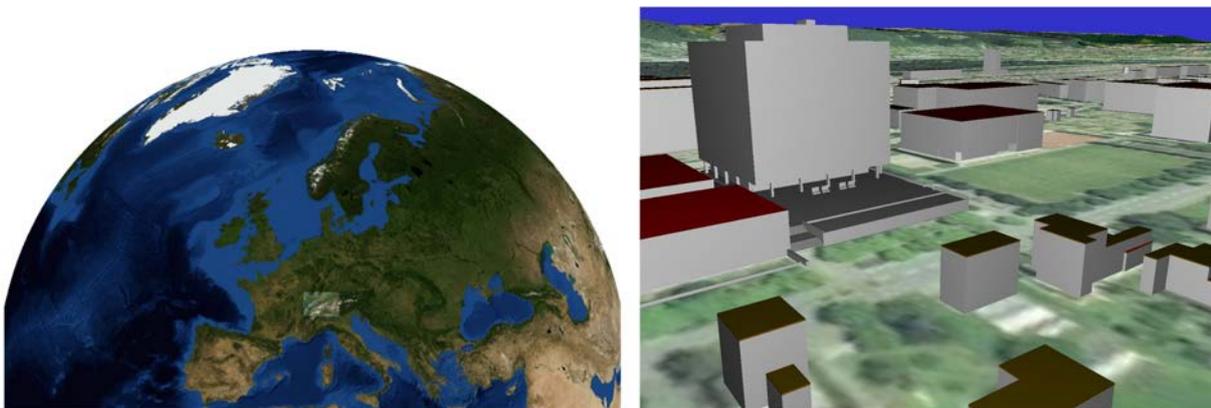


Abbildung 1-1: 3D-Geoinformationssystem auf Basis eines virtuellen Globus (links) mit integriertem digitalem 3D-Stadtmodell (rechts)

Entsprechende Bestrebungen zur Behebung dieser Defizite versprechen in Zukunft ein grosses Potential für die Umsetzung unterschiedlicher neuer professioneller 3D-Geoinformationsdienste.

Gleichzeitig haben sich über die letzten Jahre unbemannte Flugsysteme – oder synonym dazu Drohnen – stark weiterentwickelt und sind heute in unterschiedlichsten Ausprägungen und Varianten speziell auch für zivile Anwendungen auf dem Markt erhältlich. Die heutigen Flugplattformen erlauben das zusätzliche Mitführen von unterschiedlichen Sensoren die nicht zwingend für die Stabilisierung oder Navigation der Plattform selbst benötigt werden. Damit eignen sich gerade die leichtgewichtigen Mikro- und Minidrohnen als kostengünstige, schnell verfügbare und effiziente alternative Aufnahmeplattform für die Erfassung beliebiger Geodaten (Eisenbeiss, 2009) (Grenzdörffer und Engel, 2008) (Nebiker et al., 2008). Die schnelle und flexible Einsatzbereitschaft macht diese Systeme interessant für diverse Anwendungen in denen beispielsweise Echtzeit- oder Nahechtzeitbildinformationen benötigt werden. Insbesondere für die effiziente Georeferenzierung der erfassten Bilddaten sind die verfügbaren Fluglagedaten aus der Navigationseinheit von grossem Interesse. Für die Bestimmung des Fluglagezustandes kommen auf Mikro- und Minidrohnen aus Kosten- und Gewichtsgründen Navigationssensoren geringer Qualität zum Einsatz.



Abbildung 1-2: Mikrodrohne (links) mit Bodenkontrollstation (rechts)

Die Nutzung von Mikro- und Minidrohnen zur Erfassung beliebiger Echtzeit-Geodaten und deren anschliessende direkte Integration und Nutzbarmachung in entsprechenden 3D-Geoinformationsdiensten verspricht viele neue Anwendungsmöglichkeiten. Denkbar sind beispielsweise Infrastrukturmonitoring, Entscheidungssupport-Anwendungen nach Naturkatastrophen, mobile Verkehrsüberwachung, Waldbrandüberwachung und -bekämpfung oder Überwachungsaufgaben während Grossanlässen. Dafür können die georeferenzierten Bild- und Videodaten auf unterschiedliche Weise in einen 3D-Geoinformationsdienst integriert und den Nutzenden zur Verfügung gestellt werden. Möglich ist beispielsweise eine *Augmented Monitoring* (AM) Integration (vgl. Abbildung 1-3) in welcher die erfasste Videosequenz mit Objekten aus der virtuellen Welt überlagert wird. Weiter denkbar ist eine *Rapid Mapping* Integration bei der aus den

erfassten Einzelbildern mit Hilfe einer online Orthophotoprozessierung fortlaufend das Ortho-Texturmosaik im virtuellen Globus aktualisiert wird. Voraussetzung dafür ist eine möglichst akkurate Echtzeit-Georeferenzierung der erfassten Videosequenzen oder Einzelbilder, um diese anschliessend in einen 3D-Geoinformationssystem zu integrieren und über diesen zu nutzen. Gleichzeitig können dafür die zur Verfügung stehenden digitalen Landschaftsmodelldaten aus dem 3D-Geoinformationssystem für die Verbesserung der Georeferenzierungsgenauigkeit verwendet werden. Gerade die Realisierung der erwähnten *Augmented Reality* (AR) Lösung erfordert eine genaue Echtzeit-Registrierung der erfassten Bilddaten auf die bestehenden Geodaten.



Abbildung 1-3: Beispiel Augmented Monitoring Integration

Eine weitere interessante Anwendung ist die Nutzung der erfassten Bilddaten in Kombination mit kostengünstigen inertialen Navigationssystemen (INS) für die automatische Positions- und Ausrichtungsbestimmung der Aufnahmeplattform bezüglich einer bestehenden Geodatengrundlage. So kann dieses Konzept beispielsweise für eine INS/GNSS/Vision-basierte Positions- und Ausrichtungsbestimmung mobiler Trägerplattformen verwendet werden, welche die Navigation bezüglich bekannter in einem 3D-Stadtmodell modellierter Objekte ermöglicht. Besonders interessant ist dieser Lösungsansatz für die Realisierung einer Navigationslösung in Strassenschluchten bei grossen GNSS-Signalabschattungen (Globale Navigationssatellitensysteme) oder für Innenraumanwendungen.

1.2 Ziele und Abgrenzung

Mit der vorliegenden Arbeit sollten Methoden und Verfahren entworfen und in einem Prototypsystem implementiert und untersucht werden, die es ermöglichen, mobil erfasste Videosequenzen sowie Einzelbilder mit möglichst hoher Genauigkeit in Echtzeit zu georeferenzieren und gegebenenfalls auf bestehende 3D-Landschafts- und Stadtmodelle zu registrieren. Dabei wird angenommen, dass die Bilderfassungsplattform im Allgemeinen mit leichtgewichtigen und kostengünstigen Positionierungs- und Orientierungssensoren geringer Qualität ausgerüstet ist. Zudem wird vorausgesetzt, dass eine Schnittstelle zu einem virtuellen Globus bereit steht, welche es erlaubt, bestehende 3D-Landschafts- und Stadtmodelldaten (inkl. Geometrie und Semantik) innerhalb eines definierten Sichtkegels (engl. *View Frustum*) in Echtzeit für die Registrierung sowie für eine mögliche Augmentierung zur Verfügung zu stellen.

Für die Arbeit stehen folgende Fragestellungen im Vordergrund:

- Adaption des heute in der Photogrammetrie weit verbreiteten und etablierten Ansatzes der direkten Georeferenzierung (DG) für die Georegistrierung von Videodaten unter Verwendung von INS- und GNSS-Sensoren geringer Qualität.
- Entwurf und Implementierung eines integrierten Ansatzes, welcher zusätzlich zur Verfügung stehende digitale 3D-Stadtmodelle nutzt, um die erreichbare Georegistrierungsgenauigkeit aus der direkten Georeferenzierung zu verbessern.
- Entwicklung und Untersuchung von Algorithmen, die für einzelne Schlüsselbilder (engl. *Key Frames*) einer Videosequenz die äussere Orientierung mittels indirekter Georeferenzierung zur Verfügung stellen. Für diesen Prozess werden Algorithmen benötigt, die entsprechende Objektstrukturen automatisch aus der Videosequenz extrahieren und korrespondierenden 3D-Objektgeometrien aus dem vorliegenden 3D-Stadtmodell zuordnen.
- Genauigkeits- und Potentialabschätzung einer rückwärtsschnitt- vs. bündelblock-basierten indirekten äusseren Orientierungsschätzung.
- Entwurf und Implementierung eines INS/GNSS/Vision-Datenintegrationsansatzes, welcher bestehende INS/GNSS-Daten geringer Qualität mit photogrammetrisch indirekt geschätzten äusseren Orientierungen fusioniert.

Für die Video- und Bilddatenerfassung kommen in dieser Arbeit als Aufnahmeplattformen Mini- und Mikrodrohnen zum Einsatz. Für die Untersuchungen zum integrierten Ansatz soll der an der Fachhochschule Nordwestschweiz entwickelte virtuelle Globus i3D eingesetzt werden. Diese Technologie erlaubt einen effizienten Datenzugriff auf bestehende 3D-Stadtmodellgeometrien.

Die definierten Zielsetzungen unterscheiden sich wie folgt von ähnlichen Fragestellungen in benachbarten Forschungsgebieten. Im Gegensatz zu vielen *Augmented Reality* Lösungsansätzen die meist in einem lokalen Koordinatensystem arbeiten, soll in dieser Arbeit das erfasste Video zu einem übergeordneten geodätischen Referenzsystem in Bezug gesetzt werden und damit die Integration realer wie auch künstlicher Geoobjekte mit bestmöglicher Georegistrierungsgenauigkeit ermöglichen. Ähnliche Fragestellungen und Lösungsansätze werden im Forschungsgebiet *Simultaneous Localisation and Mapping* (SLAM) bearbeitet und diskutiert (Leonard und Durrant-Whyte, 1991). Vielfach werden hier unterschiedliche Kameraaufnahmekonfigurationen oder Raumvektorsensoren verwendet um kontinuierlich ein lokales Modell der erfassten Umgebung aufzubauen, welches anschliessend wieder für die Stützung der Navigation herangezogen wird. In diesem Lösungsansatz wird demzufolge für die Stützung der Navigationslösung nicht auf ein in einem übergeordneten geodätischen Referenzsystem bereits vorliegendes Stadtmodell zurückgegriffen, sondern kartiert die Umgebung in einem lokalen Referenzsystem fortlaufend neu. Eine ähnliche Vorgehensweise verfolgen auch die *Structure from Motion* (SfM) Ansätze aus der Computer Vision, wobei dort im Unterschied zu SLAM vielfach eine Monokamera-konfiguration verwendet wird.

Zusammenfassend darf festgehalten werden, dass diese Arbeit im Gegensatz zu vielen ähnlichen Ansätzen versucht, die Beziehung zwischen erfassten Bilddaten und einem bereits vorhandenen digitalen 3D-Geodatensatz herzustellen und diese anschliessend beispielsweise nutzt, um die Navigation bezüglich des übergeordneten geodätischen Referenzsystems zu stützen.

1.3 Überblick über die Arbeit

In Kapitel 2 werden zuerst vergleichbare Lösungsansätze für die definierte Problemstellung im Detail vorgestellt und diskutiert. Dabei werden vor allem Entwicklungen und Ansätze aus der nicht photogrammetrischen Sichtweise präsentiert. In einem zweiten Teil werden die benötigten theoretischen Grundlagen für die Umsetzung der erarbeiteten Methode besprochen. In diesem Zusammenhang wird auf bekannte Zuordnungsalgorithmen und -strategien, Datenfusionsansätze sowie auf bekannte Georeferenzierungsansätze und die Online-Bündelblockausgleichung eingegangen. Kapitel 3 bespricht anschliessend die in dieser Arbeit vorgeschlagene Methode. Zuerst wird das vielseitig einsetzbare Lösungskonzept dargelegt. Folgend werden die beiden aufeinander aufbauenden Georegistrierungsansätze a) Adaption der direkten Georeferenzierung auf diese Problemstellung und b) der entworfene integrierte Ansatz im Detail eingeführt. In diesem Zusammenhang werden der dafür benötigte Zuordnungsalgorithmus zwischen Einzelbildern und bestehenden 3D-Stadtmodellen, die unabhängige Orientierungsschätzung sowie die vorgeschlagene INS/GNSS/Vision-Datenintegration besprochen. Kapitel 4 vermittelt anschliessend einen Überblick über

das Programmsystem UAVision und die implementierte Video-Georeferenzierungs-Prototyplösung. Das Kapitel 5 'Ergebnisse und Testdaten' zeigt die Leistungsfähigkeit der hier vorgeschlagenen Methode auf und beurteilt diese. Dafür werden unterschiedliche mit Mini- und Mikrodrohnen erfasste Videosequenzen ausgewertet und analysiert. Das abschliessende Kapitel fasst die Arbeit zusammen und gibt einen Ausblick auf mögliche Verbesserungen und weitere offene Forschungsfragen.

2 Konzepte und Grundlagen

Dieses Kapitel vermittelt als erstes einen Überblick über verwandte Lösungsansätze und -strategien und versucht diese in Bezug auf diese Arbeit einzustufen und zu beurteilen. Anschliessend werden die relevanten theoretischen Grundlagen, mögliche Lösungsansätze und Algorithmen diskutiert, welche benötigt werden, um das in Kapitel 3 vorgeschlagene Konzept zu implementieren und umzusetzen. Abschliessend werden weitere für die Umsetzung benötigte Basistechnologien kurz eingeführt und entsprechende Anforderungen an diese formuliert.

2.1 Konzepte zur Video-Georeferenzierung

Wie in der Einleitung bereits erwähnt, existieren unterschiedlichste Konzepte und Lösungsansätze aus verschiedenen Fachgebieten, die ähnliche Ziele verfolgen und mögliche Lösungsvarianten aufzeigen. Die nachfolgenden Abschnitte gliedern die unterschiedlichen Konzepte nach deren Anwendungszweck und Fachgebiet. Die einzelnen Methoden stehen meist nicht in direktem Zusammenhang mit den Zielsetzungen dieser Arbeit, zeigen aber mögliche Lösungen einiger Teilprobleme auf. Sämtliche Ansätze beschäftigen sich zumindest damit, wie die erfasste Videosequenz oder entsprechende Einzelbilder in Bezug zu einer digitalen Modellbeschreibung der realen Umgebung gesetzt werden können. Am Ende werden die einzelnen Methoden zur Gewinnung eines Überblicks miteinander verglichen und hinsichtlich der Eignung zur Problemlösung beurteilt.

2.1.1 Direkte und integrierte Georeferenzierung in der Photogrammetrie

Der Ansatz der direkten sowie der integrierten Georeferenzierung, wie er heute für die Erfassung von Geodaten aus satelliten-, flugzeug- oder fahrzeuggestützten Aufnahmeplattformen verwendet wird, ermöglicht auch die Georegistrierung mobil erfasster Videodaten. Diese beiden Auswertemethoden wurden im Speziellen für die direkte Bestimmung der Bildorientierung photogrammetrischer Aufnahmen mit Hilfe von INS/GNSS-Navigationssensoren hoher Qualität entwickelt. Die dabei erreichbare Georeferenzierungsgenauigkeit hängt im Wesentlichen von der Qualität der eingesetzten INS/GNSS-Sensoren ab. Die entwickelten Datenverarbeitungskonzepte sind auf die Güte dieser Sensorkategorie ausgelegt und können nicht ohne Adaption und Einschränkungen auf INS/GNSS-basierte Navigationslösungen geringer Qualität übertragen werden. Eine Gegenüberstellung sowie die theoretischen Grundlagen dieser beiden Ansätze werden im Detail im Abschnitt 2.3 dieses Kapitels besprochen.

Unter dem Begriff der direkten Georeferenzierung wird das direkte Messen der Elemente der äusseren Orientierung für einen beliebigen Geosensor verstanden. Dabei wird die gesuchte äussere Orientierung mit Hilfe von INS- und GNSS-Empfänger direkt bestimmt und nicht mittels indirekter Georeferenzierung (vgl. Abschnitt 2.3.1) durch

eine konventionelle Aerotriangulation geschätzt. Erste konzeptionelle Überlegungen zur direkten Georeferenzierung gehen auf Schwarz et al. (1984) zurück. Nachfolgend wurde diese Technologie hauptsächlich am Department of Geomatics Engineering der University of Calgary weiterentwickelt (Schwarz et al., 1993) (Schwarz, 1995) (Skaloud et al., 1996) (Skaloud, 1999a). Weitere wesentliche Untersuchungen wurden am Institut für Photogrammetrie der Universität Stuttgart (Cramer, 2001a) (Haala, 2005) geleistet, die zur heutigen Etablierung der Messmethode beigetragen haben. Parallel dazu hat sich die direkte Georeferenzierung mit der Einführung von flugzeuggetragenen digitalen Zeilenkameras und LiDAR-Sensoren entwickelt (Müller, 1991) (Sandau und Eckert, 1996) (Wewel et al., 1998), welche für die Georeferenzierung der erfassten Daten die direkte kontinuierliche Messung der Parameter der äusseren Orientierung voraussetzen. Eine Zusammenstellung der Entwicklungen der direkten Georeferenzierung ist in Mostafa and Hutton (2005) zu finden. Zur Messdatenintegration aus INS und GNSS für die Bestimmung der äusseren Orientierung wird heute ein *loosely-* oder *tightly coupled* Kalman-Filter (KF) eingesetzt (vgl. Abschnitt 2.4). Ein Überblick über unterschiedliche verwendete Kalman-Filter-Integrationsstrategien in der Luftbildphotogrammetrie ist beispielsweise in Haala (2005) oder Ellum und El-Sheimy (2006) zu finden. Im Unterschied zum Ansatz der direkten Georeferenzierung werden beim integrierten Ansatz zusätzlich zu den direkt durch INS und GNSS gemessenen äusseren Orientierungselementen Bildmessungen auf bekannte Passpunkte zur Schätzung und Korrektur einer allfällig enthaltenen Restsystematik verwendet (Cramer, 2001a).

2.1.2 Video Registration

Das Fachgebiet *Video Registration* hat sich in den letzten zehn Jahren im Umfeld der Computer Vision entwickelt und beschäftigt sich mit ähnlichen Fragen wie diese Arbeit. Hauptaufgabe bildet die Registrierung einer mobil erfassten Videosequenz auf eine bestehende Geodatengrundlage. Dabei wird vielfach auch mit unbemannten Flugsystemen in Kombination mit Navigationssensoren geringer Qualität gearbeitet. Viele in diesem Umfeld publizierte Ansätze verwenden die zur Verfügung stehenden Fluglagedaten der Video-Trägerplattform als Näherungswerte für die anschliessende Fein-Georegistrierung. Eine Zusammenstellung bestehender Ansätze ist in Sheikh et al. (2003) dargelegt. Grob können die publizierten Ansätze wie folgt unterschieden werden:

Intensitätbasierte Georegistrierung

Diese Lösungsansätze versuchen die Videosequenz auf ein bestehendes Orthophotomosaik zu georegistrieren. Bei dieser Klasse erfolgt die Georegistrierung durch eine merkmalsbasierte Zuordnung (engl. *Feature Based Matching*) zwischen den Einzelbildern der Videosequenz und dem orthorektifizierten Referenzbild. Unterschiedliche Implementierungen dieses Ansatzes werden beispielsweise in Sheikh et al. (2005), Cannata et al. (2000) oder Kumar et al. (1998) vorgestellt. Gerade für die Georeferenzierung mobil erfasster Videosequenzen im offenen Gelände ist dieser Lösungsansatz vielversprechend. Jedoch werden auch hier einige Schwierigkeiten

erwähnt. So ist beispielsweise eine grosse zeitliche Differenz zwischen verfügbaren Referenzdaten und erfassten Videodaten für die automatische texturbasierte Zuordnung unter Umständen problematisch. Viele der vorgestellten Implementierungen sind zudem nicht vollständig Echtzeitfähig.

Höhenmodellbasierte Georegistrierung

Bei diesem Lösungsansatz wird versucht, die Videosequenz auf ein bestehendes Höhenmodell zu registrieren. Dabei wird aus nachfolgenden Bildern der Videosequenz vorgängig ein Höhenmodell abgeleitet. Anschliessend wird versucht, dieses Modell auf das bestehende Referenzhöhenmodell aufgrund eindeutiger Oberflächenstrukturen zuzuordnen. Dieser Ansatz ist als mögliches Gedankenmodell zu verstehen, eine entsprechende praktische Umsetzung ist dem Autor bis anhin nicht bekannt.

3D-Modell-basierte Georegistrierung

Der 3D-Modell-basierte Ansatz verwendet als Referenz 3D-Vektordaten bestehender digitaler Stadtmodelle und versucht Objektstrukturen aus dem Vektormodell in der Videosequenz zu erkennen und entsprechend zuzuordnen. Eine Implementierung dieses Ansatzes wird in Matai et al. (2003) aufgezeigt und sollte besonders für Anwendungen im Siedlungsgebiet im Vergleich zur texturbasierten Methode bessere Resultate liefern. Die Methode von Matai extrahiert mit Hilfe eines Gradientenfilters Kanten von Objektstrukturen im Videobild. Die eigentliche Registrierung zwischen extrahierten Kanten und dem 3D-Stadtmodell wird mit Hilfe einer Energiefunktional-Optimierung erreicht (vgl. Abschnitt 2.2.1). Dabei definiert sich das Energiefunktional aus der summierten Gradientenenergie der mit den aktuellen äusseren Orientierungsparametern ins Bild zurückprojizierten 3D-Modellkanten. Nachteil dieses Optimierungsansatzes ist der sehr hohe Rechenaufwand, was die Nutzung in einer Echtzeitanwendung zurzeit verunmöglicht. Weiter werden sehr gute Näherungswerte für Aufnahmeposition- und Orientierung aus den Fluglagedaten benötigt, damit der Optimierungsalgorithmus ins gesuchte lokale Optimum konvergiert.

2.1.3 Objektrekonstruktion aus Bildsequenzen

Unter dem Begriff *Structure from Motion* (SfM) versteht man die automatische 3D-Rekonstruktion eines mittels bewegter Videokamera erfassten Objektes. Entsprechende Methoden zur Lösung dieser Fragestellung wurden in der Computer Vision entwickelt und finden heute beispielsweise vermehrt Anwendung in der Nahbereichsphotogrammetrie. Grundsätzlich rekonstruiert diese Methode aus den einzelnen Bildern der Sequenz das erfasste Objekt dreidimensional und bestimmt gleichzeitig die Aufnahmeposition und -ausrichtung der Einzelbilder in Bezug zu diesem Objekt. Genau diese vollautomatisch bestimmte Beziehung zwischen rekonstruierter 3D-Objektgeometrie und des Kameraaufnahmepfades (Äussere Orientierung der Aufnahmegeometrie) ist für diese Arbeit von Interesse. Vielfach wird die SfM Problemstellung mit Hilfe der ebenfalls in der Computer Vision entwickelten Algorithmen der *Multiple View Geometry* gelöst. Diese Theorie beschreibt in der projektiven Geometrie

die relative gegenseitige Beziehung mehrerer Bildaufnahmen unter gleichzeitigem Einbezug der erfassten Objektgeometrie. Eine detaillierte Besprechung dieser Algorithmen ist nicht Bestandteil dieser Arbeit und kann der einschlägigen Fachliteratur entnommen werden. Eine gute Einführung in die *Multiple View Geometry* im Speziellen für Photogrammeter wird in Ressl (2003) oder Rodehorst (2004) vermittelt. Ansonsten sei auf die beiden Referenzwerke von Hartley und Zisserman (2003) und Faugeras (1993) verwiesen.

2.1.3.1 Structure from Motion

Abbildung 2-1 zeigt die grundlegenden Auswerteschritte wie aus einer Videosequenz oder aus Einzelbildern automatisch das aufgenommene Objekt und die Aufnahmegeometrie rekonstruiert wird. Eine vertiefte Darstellung dieser Methode ist in Pollefeys et al. (1999) zu finden. Zu erwähnen bleibt hier, dass es auch möglich ist, mit Videosequenzen zu arbeiten, bei denen die Parameter der inneren Orientierung vorgängig unbekannt oder variabel sind wie die Veröffentlichungen von Pollefeys (1999) und Pollefeys und Van Gool (1999) aufzeigen. Die höchste geometrische Genauigkeit, wie es meist photogrammetrische Anwendungen erfordern, wird durch eine anschließende nachträgliche Bündelblockausgleichung (vgl. Abschnitt 2.3.1.3) über sämtliche Aufnahmen der Bildsequenz erzielt. Dafür wird die mit der *Multiple View Geometry* Methode automatisch rekonstruierte Aufnahme- und Objektgeometrie als Näherungslösung in die Bündelblockausgleichung eingeführt, wie dies beispielsweise Rodehorst (2004) aufgezeigt.

2.1.3.2 Augmented Reality Anwendungen

Grundlage für viele 'Erweiterte Realität' Anwendungen (engl. *Augmented Reality*) bildet ein digitales Abbild der realen Umgebung in welche das virtuelle Objekt hineingesetzt werden soll. Zusätzlich muss der Kamerapfad der Videosequenz, in welche das virtuelle Objekt integriert werden soll, bezüglich der digitalen Repräsentation der realen Umgebung als bekannt vorliegen. Es können nun folgende zwei Fälle unterschieden werden:

- a. Es liegt vorgängig kein Modell der realen Umgebung vor: Bei dieser Ausgangslage kann die Rekonstruktion der realen Umgebung und die Lage und Ausrichtung der Aufnahmegeometrie mittels SfM-Methode generiert werden (Cornelis et al., 2001).
- b. Es liegt eine digitale Repräsentation der Umgebung vor: In diesem Fall muss lediglich die Lage und Ausrichtung der Aufnahmegeometrie bezüglich der vorliegenden digitalen Umgebungsrepräsentation bestimmt werden. Dafür können die diskutierten Verfahren und Ansätze in diesem Abschnitt verwendet werden.

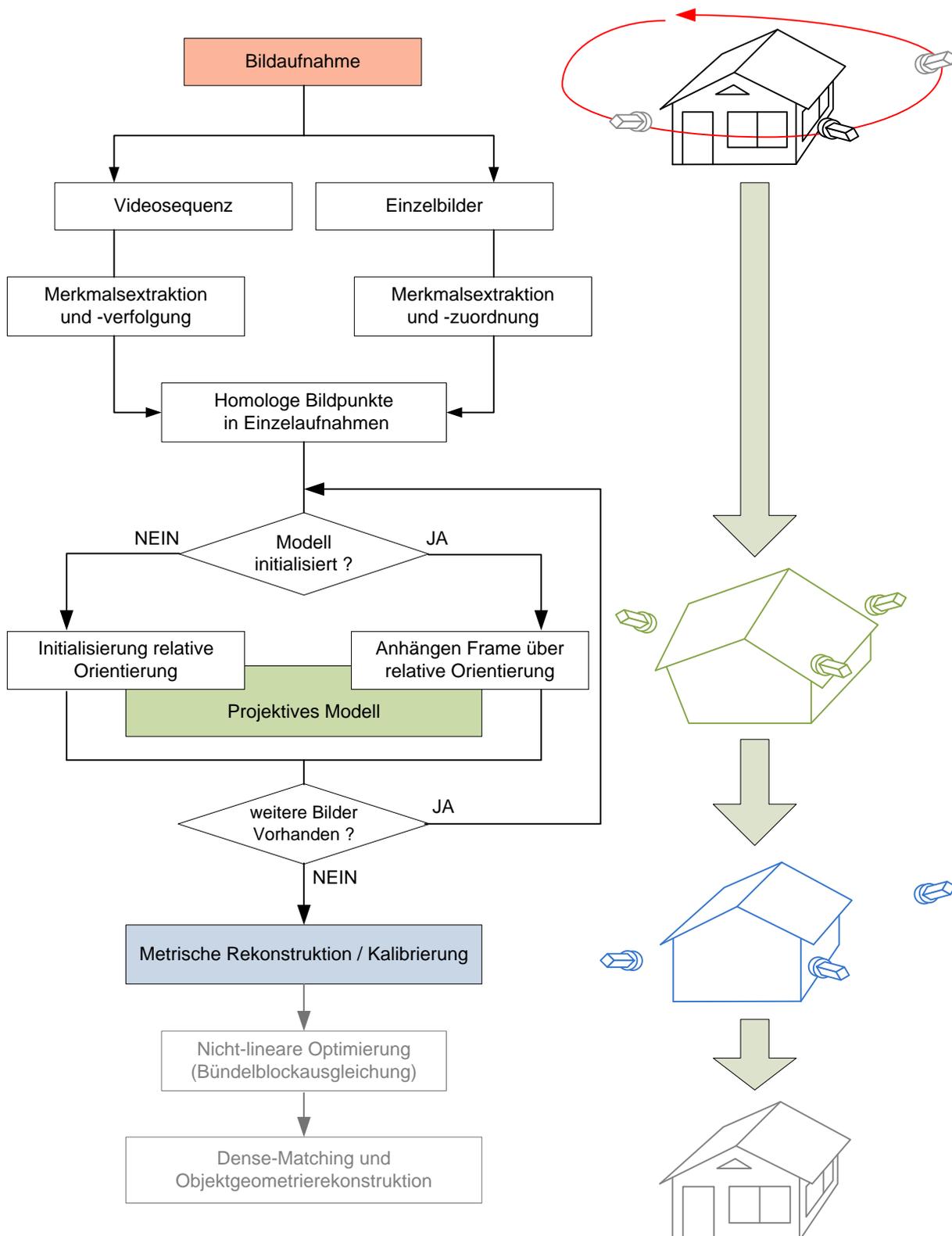


Abbildung 2-1: Structure from Motion Datenverarbeitung

Basierend auf der nun zur Verfügung stehenden Information kann mit Hilfe der Computergraphik das virtuelle Objekt entsprechend in der Videosequenz überlagert werden. Beispielsweise setzt die Filmindustrie für ihre Produktionen vielfach diese Technologie ein und spricht in diesem Zusammenhang von *Matchmoving* (Dobbert,

2005). Zu diesem Zweck existieren verschiedene kommerzielle Softwareprodukte, die SfM-Algorithmen zur Lösung der Problemstellung implementieren.

2.1.4 Simultaneous Localisation and Mapping

Unter dem Begriff *Simultaneous Localisation and Mapping* versteht man allgemein das Lokalisieren der Aufnahmesensoren und die gleichzeitige Kartierung der Umgebung mit diesen. Dabei wird für die Lokalisierung wiederum die vorgängig kartierte Umgebung eingesetzt. Das Aufnahmesystem einer SLAM-Anwendung besteht meist aus unterschiedlichen Sensoren wie Stereo-, Monokameras, Rangekameras oder Vektormesssystemen – gegebenenfalls ergänzt durch ein INS oder einen GNSS-Empfänger. Die Problemformulierung wie auch entsprechende Lösungsansätze haben sich über die letzten zwei Jahrzehnte hauptsächlich innerhalb des Fachgebiets der Robotik entwickelt. In Durrant-Whyte et al. (1996) wurde die Fragestellung unter dem Begriff SLAM erstmals definiert und eingeführt. In der anschließenden Dekade erfolgten viele Arbeiten und Publikationen zu diesem Thema. SLAM wird heute beispielsweise für die Navigation von Robotersystemen für Innenraum-, Aussenraum-, Unterwasser- oder luftgestützte Anwendungen eingesetzt. Nachfolgende Abbildung 2-2 verdeutlicht das SLAM-Prinzip.

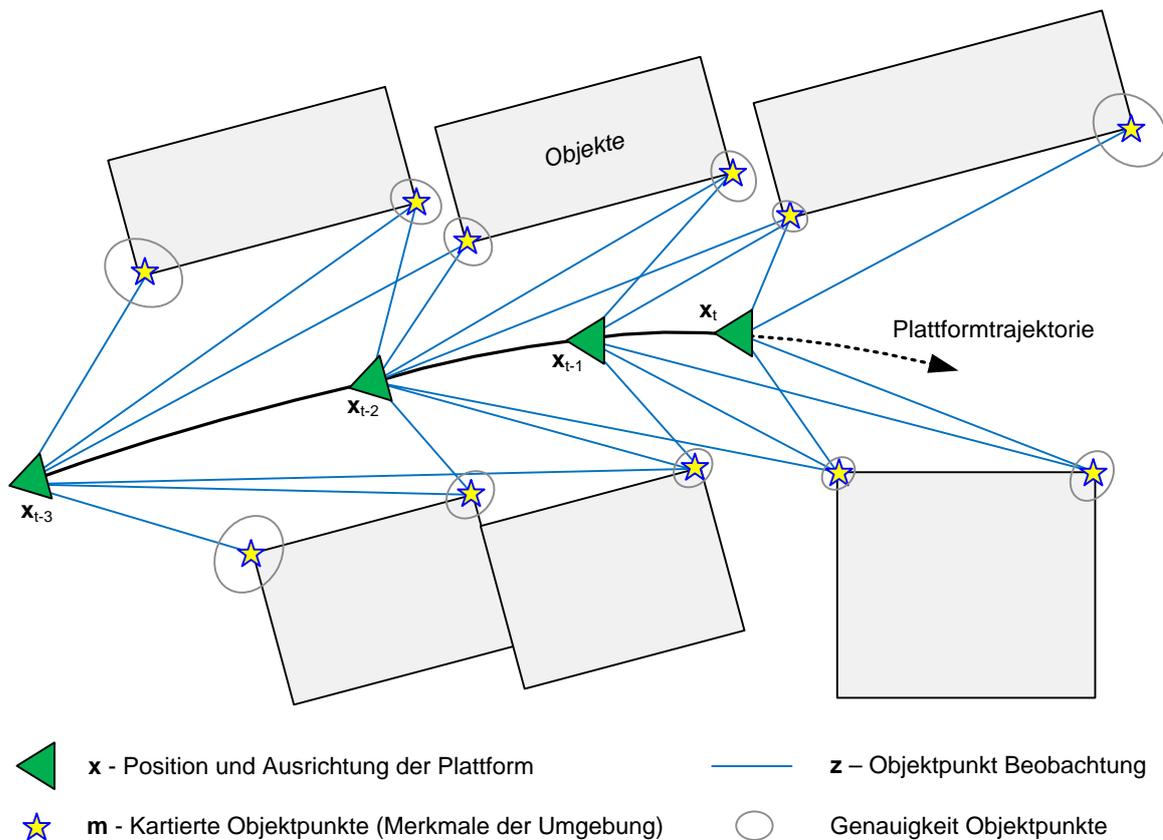


Abbildung 2-2: SLAM Prinzip

In SLAM werden der aktuelle Zustand der Aufnahmeplattform \mathbf{x}_t auf der einen Seite sowie die Position der kartierten Merkmale \mathbf{m}_i auf der anderen Seite mittels einer Wahrscheinlichkeitsverteilung beschrieben. Dabei wird die Aufnahmeplattform mit einer angepassten Systemdynamik beschrieben, aus welcher meist die Position und Orientierung bezüglich der kartierten Merkmale abgeleitet werden kann. Demgegenüber wird die Umgebung mit Hilfe statischer Merkmale meist in Form einer 3D-Position im Zustandsvektor abgebildet. Für die Lösung der SLAM-Problemstellung wird im Allgemeinen folgender rekursiver Zweischritt-Algorithmus verwendet:

- a. Prädiktion des Systemzustandes: Mit Hilfe der bekannten Steuergrößen des Robotersystems (Aufnahmeplattform) und des daraus ableitbaren Bewegungsmodells lässt sich aus dem letzten Aufnahmesystemzustand \mathbf{x}_{t-1} der aktuelle Zustand \mathbf{x}_t präzisieren.
- b. Korrektur des Systemzustandes: Mit den zum Zeitpunkt t zur Verfügung stehenden Beobachtungen \mathbf{z}_t auf bereits kartierte Merkmale lassen sich durch ein entsprechendes Beobachtungsmodell der gesamte Systemzustand aktualisieren und gegebenenfalls n neue Merkmale \mathbf{m}_n initialisieren.

Eine entsprechende Implementierung erfolgt im Allgemeinen basierend auf einem der beiden Algorithmen:

- EKF-SLAM: In dieser Variante wird für die Umsetzung auf das erweiterte Kalman-Filter (EKF - engl. *Extended Kalman Filter*) zurückgegriffen. Dieses Filter erlaubt die Beschreibung des Aufnahmesystemzustands und der Merkmale mit Hilfe einer Gauss-Wahrscheinlichkeitsverteilung. Nachteil dieser Implementierungsvariante ist der resultierende riesige Zustandsvektor, womit eine Echtzeitverarbeitung gerade für mehrere Tausend involvierte Merkmale nicht gewährleistet ist.
- FastSLAM: Diese Implementierungsvariante basiert auf einem Partikel-Filter, der für die Beschreibung des Systemzustandes eine beliebige Wahrscheinlichkeitsverteilung erlaubt. Mit Hilfe des vielfach verwendeten Rao-Blackwellisation Algorithmus lässt sich zusätzlich der riesige Zustandsvektor reduzieren, was eine Echtzeitverarbeitung mit tausenden von Merkmalen ermöglicht.

Eine weiterführende Betrachtung der Problemstellung und der grundsätzlichen Lösungsansätze ist in Durrant-Whyte und Bailey (2006) zusammengefasst. Aktuelle Trends und Algorithmen in SLAM werden beispielsweise in Bailey und Durrant-Whyte (2006) besprochen. Gerade SLAM-Anwendungen, basierend auf einem Monokamera-Aufnahmesystem, sind im Zusammenhang mit dieser Arbeit von grossem Interesse. Die entsprechende Schwierigkeit in Monokamera SLAM-Anwendungen bildet die zuverlässige Initialisierung der Merkmale aus Bildsequenzen mit kleiner Bildbasisänderung. Lösungsansätze und Implementierungen werden in den Arbeiten von Farley und Chapman (2008), Montiel et al. (2006), Davison (2003) sowie Steffen und Förstner (2008) vorgestellt.

Heute werden im Zusammenhang mit mobilen Kartierungs- und Messfahrzeugen (MMS - engl. *Mobile Mapping Systems*), Innenraum-Navigationslösungen oder autonomen Flugsystemen in der Geomatik solche Methoden aufgegriffen und adaptiert. Im Gegenzug bieten gerade Algorithmen zur Bildorientierung aus der Photogrammetrie

interessante Lösungsansätze insbesondere für bildbasierte SLAM-Anwendungen. So wird heute in der Robotik unter dem Begriff *Information Filters* (Thrun et al., 2005) eine weitere Implementierungsvariante verstanden, bei welcher die Aufnahmetrajektorie sowie die Merkmale mit Hilfe einer online Bündelblockausgleichung bestimmt werden (vgl. Abschnitt 2.3.1.3). Einen weiteren interessanten Lösungsansatz zeigt die Arbeit von Bayoud (2006) auf. Bayoud verwendet einen sich fortlaufend abwechselnden Prozess zwischen Vorwärtseinschneiden für die Kartierung der Umgebung und Rückwärtseinschneiden zur Lokalisierung der Aufnahmesensoren. Dabei behandelt er im Unterschied zum klassischen SLAM-Ansatz die kartierten Punkte nicht stochastisch und damit ausserhalb des Zustandsvektors. Eine weitere sehr interessante Arbeit versucht mit Hilfe der klassischen SLAM-Methode die Navigationslösung einer Mikrodrohne mit Hilfe des auf der Plattform erfassten Videos zu stützen (Steffen und Förstner, 2008).

2.1.5 Zusammenstellung und Beurteilung

Wie in Tabelle 2-1 ersichtlich ist, unterscheiden sich der photogrammetrische Ansatz und die Konzepte aus dem Fachgebiet *Video Registration* nur marginal. Entscheidend ist jedoch, dass der photogrammetrische Lösungsansatz und die dafür entwickelten Auswerteverfahren und theoretischen Grundlagen davon ausgehen, dass für die Bestimmung der Flugtrajektorie Sensoren hoher Qualität verwendet werden. Demgegenüber gehen die *Video Registration* Ansätze davon aus, dass die zur Verfügung stehenden Navigationsdaten aus Sensoren geringer Qualität geschätzt und demzufolge lediglich als Näherungswerte im eigentlichen Bilddatenregistrierungsprozess verwendet werden. Gemeinsam ist jedoch beiden Ansätzen, dass sie die erfassten Bilddaten bezüglich eines geodätischen Bezugssystems beziehungsweise einer existierenden Geodatenbasis referenzieren. Im Gegensatz dazu werden beim SfM- und SLAM-Ansatz die Ausrichtungs- und Positionsdaten des Aufnahmesensors sowie die rekonstruierte Umgebung grundsätzlich in einem lokalen Bezugssystem bestimmt und verwenden daher meist keine bestehende Geodatenbestände. Jedoch zeigen gerade diese beiden Ansätze interessante Konzepte und Algorithmen auf, welche die vollautomatische Positionierung und Ausrichtung von Bildsensoren ermöglichen. Für die Umsetzung der definierten Aufgabenstellung wird für die Video-Georegistrierung ein zweistufiger Ansatz vorgeschlagen, welcher für die Grobregistrierung die direkte Georeferenzierung nutzt und darauf aufbauenden eine Feinregistrierung mit Hilfe einer integrierten Georeferenzierung durchführt. Die Feinregistrierung integriert Konzepte und Lösungsansätze aus SfM und SLAM und ermöglicht insbesondere die Registrierung der erfassten Bilddaten auf eine bestehende Geodatenbasis.

	Direkte- und integrierte Georeferenzierung	Video Registration	Structure from Motion	SLAM
Sensoren	INS $x^{(Q+)}$	$x^{(Q-)}$		$(x)^{(Q-)}$
	GNSS $x^{(Q+)}$	$x^{(Q-)}$		$(x)^{(Q-)}$
	Bildkamera x	x	x	x
	weitere			(x)
Registrierung bez. lokalem Koordinatensystem (relativ)			x	x
Registrierung bez. geodät. Referenzsystem (absolut)	x	x		
Kartierung aus Sensordaten	x		x	x
Prozessierung	Offline	On-/Offline	Offline	Online

$^{(Q+)}$ - Sensor hoher Qualität, $^{(Q-)}$ - Sensor geringer Qualität

Tabelle 2-1: Lösungsansätze und -strategien im Vergleich

2.2 Grundlagen zur Bild-zu-3D-Stadtmodell-Zuordnung

Unter dem englischen Begriff *Matching*, der heute vielfach auch im Deutschen Verwendung findet, wird nach Heipke (1996) Zuordnung oder auch Korrelation verstanden. Ziel eines Zuordnungsverfahrens oder Matchings ist das Auffinden zweier oder mehrerer korrespondierender Beschreibungen desselben Objekts. An dieser Stelle sei erwähnt, dass sich über die letzten Dekaden eine grosse Anzahl an unterschiedlichen Zuordnungsverfahren und -strategien entwickelt und etabliert hat. Es ist aus verschiedenen Gründen eine schwierige Aufgabe eine allgemeingültige Einteilung oder Klassifizierung der Methoden zu finden. Zum einen wurden die einzelnen Verfahren in verschiedensten Fachgebieten entwickelt und zum anderen verfolgen sie unterschiedliche Zielsetzungen aus verschiedenen Anwendungsbereichen. In diesem Zusammenhang werden die Ansätze nach unterschiedlichen Anforderungen und Sichtweisen beurteilt und entsprechend klassiert. In Baltsavias (1991) ist eine Klassifizierung von Verfahren zu finden, die im Wesentlichen Bild-zu-Bild-Zuordnungsprobleme lösen. Der nachfolgende Abschnitt vermittelt einen Überblick über Zuordnungsverfahren in Bezug auf die Fragestellung in dieser Arbeit. Im Abschnitt 2.2.2 werden anschliessend für diese Arbeit relevante Bild-zu-Modell-Zuordnungsverfahren behandelt.

2.2.1 Zuordnungsverfahren

Die Vielzahl der Zuordnungsverfahren können nach unterschiedlichen Kriterien unterteilt werden. Nachfolgend wird eine Klassifizierung vorgeschlagen, welche die Verfahren nach den Gesichtspunkten Zuordnungsebene, Zuordnungsmethode und Optimierungsansatz gliedert.

Für die Beschreibung eines Zuordnungsproblems wird die in Shapiro and Haralick (1981) eingeführte Notation verwendet. In diesem Kontext beschreibt die Terminologie 'Zuordnung' das Auffinden einer Abbildungsfunktion h , die eine optimale Übereinstimmung der Modellbeschreibung m^A und der korrespondierenden Modellbeschreibung m^B beschreibt.

$$m^A = h \circ m^B \quad (1)$$

Im Allgemeinen kann eine Modellbeschreibung m^k abhängig vom Zuordnungsverfahren durch unterschiedliche Objektausprägungsformen aufgefasst werden. Im Speziellen bei merkmalsbasierten Zuordnungsverfahren kann grundsätzlich die Modellbeschreibung $m^k = (P)$ durch die Menge $P = \{p_1^k, p_2^k, \dots, p_n^k\}$ bestehend aus den Primitiven p_i^k gebildet werden. Dabei wird eine einzelne Primitive p_i^k aus der Modellbeschreibung k mit Hilfe

eines Merkmalsvektors $p_i^k = (v_1, v_2, \dots, v_o)^T$ durch dessen Eigenschaften v_i repräsentiert, welcher auch als Merkmalsdeskriptor der Primitive p_i^k bezeichnet wird. Weiterführende merkmalsbasierte Zuordnungsverfahren verwenden zur Beschreibung eines Modells zusätzlich Beziehungen zwischen den Primitiven der Menge P . Für diesen Fall wird die Modellbeschreibung $m^k = (P, R)$ durch die Menge $R = \{r_1^t, r_2^t, \dots, r_l^t\}$ ergänzt, welche sich aus l Relationen r_i^t unterschiedlicher Beziehungstypen t zusammensetzt. Dabei bildet sich ein einzelner Beziehungstyp durch die Funktion $r_i^t = f^t(p_i^k, p_j^k)$ abhängig von zwei Primitiven der Menge P .

Die Abbildungsfunktion h kann abhängig vom Zuordnungsverfahren sowie von der Repräsentation der Modellbeschreibung auf unterschiedliche Art interpretiert werden. Einerseits kann h als Vorschrift aufgefasst werden, welche einzelne Primitiven p_i^A der Modellbeschreibung m^A auf die Primitiven p_i^B der Modellbeschreibung m^B abbildet. Andererseits kann darunter auch der mathematische funktionale Zusammenhang $m^A = h(m^B)$ zwischen zwei Modellbeschreibungen verstanden werden.

2.2.1.1 Zuordnungsebenen

Die Zuordnungsebene definiert nun den Beziehungstyp zwischen den beiden Modellbeschreibungen und bestimmt somit den Abbildungsfunktionstyp h . Folgende Zuordnungsebenen können unterschieden werden:

- Bild-zu-Bild-Zuordnung (2D-2D):
Unter diese Zuordnungsebene fallen Verfahren, die Objekte aus einem Bild korrespondierenden Objekten im anderen Bild zuordnen.

$$m^{img} = h_{2D-2D} \circ m^{img} \quad (2)$$

Die Objekte werden hier in beiden Modellbeschreibungen m^{img} meist mit Bildausschnitten oder einem Merkmal aus dem jeweiligen Bild beschrieben. In diese Zuordnungsebene fallen die meisten Verfahren welche in der Photogrammetrie oder der Computer Vision verwendet werden. Allgemein können darunter Zuordnungsverfahren verstanden werden, die 2D-Modellbeschreibungen möglichst optimal aufeinander abbilden.

- Bild-zu-Modell-Zuordnung (2D-3D):
Verfahren dieser Zuordnungsebene suchen eine optimale Abbildung zwischen einer Modellbeschreibung eines Bildes m^{img} und einer 3D-Objektrepräsentation m^{mo} .

$$m^{img} = h_{2D-3D} \circ m^{mo} \quad (3)$$

Diese Bild-zu-Modell-Zuordnung muss nicht zwingend direkt erfolgen, sondern kann auch durch folgende zwei Schritte gelöst werden.

$$m^{img} = h_{2D-2D} \circ (h_{2D-3D}^{ini} \circ m^{mo}) \quad (4)$$

In einem ersten Teilschritt wird über eine Initiale-Abbildungsvorschrift h^{ini} die 3D-Modellbeschreibung m^{mo} in eine 2D-Modellbeschreibung abgebildet. In einem zweiten Teilschritt erfolgt nun eine Bild-zu-Bild-Zuordnung zwischen der Modellbeschreibung des Bildes m^{img} und der abgeleiteten 2D-Modellbeschreibung, woraus die gesuchte Abbildungsvorschrift h_{2D-3D} resultiert.

- **Modell-zu-Modell-Zuordnung (3D-3D):**
Diese Zuordnungsebene behandelt Verfahren die zwei 3D-Modellbeschreibung m^{mo} möglichst optimal aufeinander abbilden.

$$m^{mo} = h_{3D-3D} \circ m^{mo} \quad (5)$$

Diese Verfahren sind speziell für Anwendungen interessant, in welchen eine lokale 3D-Modellbeschreibung oder ein lokaler Koordinatenrahmen auf eine globale 3D-Modellbeschreibung oder einen geodätischen Referenzrahmen abgebildet wird.

2.2.1.2 Zuordnungsmethoden

Diese Betrachtungsebene identifiziert grundlegende Zuordnungsmethoden aus technischer Sichtweise aus der grossen Anzahl von Verfahren und fasst diese kurz zusammen. Die nachfolgend vorgestellten Methoden können für Problemstellungen auf unterschiedlichen Zuordnungsebenen eingesetzt werden und lassen sich vielfach mit Hilfe verschiedener Optimierungsalgorithmen lösen.

- **Flächenbasierte Zuordnung (ABM - engl. *Area Based Matching*):**
Diese Zuordnungsmethode kann ausschliesslich für ein Matching zwischen Merkmalen zweier Rasterbildrepräsentationen¹ verwendet werden. Als Objekt wird ein geeigneter Bildausschnitt aus dem Originalbild gewählt, das der korrespondierenden Bildstelle im Suchbild zugeordnet werden soll. Der Bildausschnitt des Originalbildes und des Suchbildes lassen sich durch die Grauwertfunktionen $o(x, y)$ und $s(x, y)$ beschreiben.

$$F\{ o(x,y), s(f_x(x,y), f_y(x,y)) \} \Rightarrow optimum \quad (6)$$

Die hier allgemein formulierte Problemstellung kann beispielsweise mit Hilfe einer Kleinste-Quadrate-Anpassung (LSM - engl. *Least Squares Matching*), dem Korrelationsverfahren oder dem Schwerpunktverfahren gelöst werden. Eine weiterführende Beschreibung dieser Verfahren ist in Luhmann (2003) zu finden.

- **Merkmalsbasierte Zuordnung (FBM - engl. *Feature Based Matching*):**
Die merkmalsbasierten Zuordnungsverfahren verwenden eine sogenannte Merkmalsbeschreibung oder einen Deskriptor um ein Objekt zu repräsentieren. Vielfach werden die beiden Objekte aus den zwei Modellrepräsentationen einander zugeordnet, welche die ähnlichste Merkmalsbeschreibung aufweisen. Viele FBM-Verfahren verwenden Primitiven wie Punkte, Kanten oder Flächen als Objekte und

¹ Aufgefasst als 2.5D-Grauwertfunktion

leiten einige Eigenschaften dieser Primitiven ab, welche anschliessend den dazugehörigen Merkmalsdeskriptor bilden. Meist müssen zuerst die gesuchten Merkmale mit Hilfe einer entsprechenden Methode – beispielsweise aus einem Grauwertbild – extrahiert und die Eigenschaften v_i des Deskriptors p_i^k abgeleitet werden. Für die Merkmalsextraktion werden vielfach sogenannte Interest-Operatoren verwendet. Bekannte Verfahren sind die Operatoren von Moravec (1977), Förstner und Gülch (1987), Harris und Stephens (1988) oder der SIFT-Operator von Lowe (1999). Es existieren unzählige Zuordnungsverfahren, auf einige die speziell Kanten-Primitiven verwenden, wird im Abschnitt 2.2.2 näher eingegangen.

- **Relationale Zuordnung (RM - engl. *Relational Matching*):**
Die relationalen Zuordnungsverfahren können als Untergruppe der merkmalsbasierten Zuordnungsverfahren eingestuft werden, da sich die beiden Methoden in einem relevanten Punkt unterscheiden. Ein relationales Verfahren lässt im Gegensatz zu den rein merkmalsbasierten Verfahren zusätzliche Beziehungen zwischen zwei Objekten einer Modellbeschreibung zu (Vosselman, 1992). Bei der Zuweisung werden nicht nur die Deskriptoren selbst, sondern zusätzlich die modellierten Beziehungen zwischen den Merkmalen in die Beurteilung der Ähnlichkeit miteinbezogen.
- **Energiefunktional-Optimierung (EFO - engl. *Energy Functional Optimisation*):**
Bei dieser Methode wird meist eine Zuordnung zwischen einem Bild und bestehenden Modelldaten gesucht (Bild-zu-Modell-Zuordnung). Für die Lösung dieser Problemstellung wird vielfach wie nachfolgend aufgezeigt ein Energiefunktional aufgestellt, welches die Ähnlichkeit global über die gesamte Modellbeschreibung definiert, die es zu optimieren gilt (Scholze, 2002).

$$EF(h) = \sum_{i=1}^n w_i E_i(m^A, m^B) \quad (7)$$

Darin werden sogenannte Energietherme E_i eingeführt, welche die Zuordnung mit Hilfe einer definierten Funktion zwischen den beiden Modellbeschreibungen beurteilen. Die einzelnen Energieterme können durch unterschiedliche Gewichtung w_i in die Gesamtbeurteilung einfließen. Vielfach wird die Ähnlichkeit beurteilt über die anfallenden Kosten oder über den Nutzen einer Zuordnung. In diesem Kontext wird das Energiefunktional als Kostenfunktion, die minimiert oder als Nutzenfunktion die maximiert werden soll, aufgefasst. Beispielsweise kann bei einer Bild-zu-Modell Zuordnung die Modellgeometrie mit Hilfe einer zur Verfügung stehenden initialen Abbildung h in das Bild projiziert werden, wo diese mit Hilfe des definierten Energiefunktionals beurteilt wird. In einem iterativen Prozess werden nun die Abbildungsparameter dahingehend verändert, dass sich das Energiefunktional optimiert. Vielfach wird dazu ein direktes Optimierungsverfahren eingesetzt. Der Energiefunktional-Optimierungsansatz wird ebenfalls erfolgreich für die Lösung von Bild-zu-Bild-Stereozuordnungsproblemen verwendet. Beispiel hierfür ist der Semi Global Matcher-Algorithmus, welcher die pixelweise Zuordnung durch

Minimierung einer global definierten Energiefunktion realisiert (Hirschmüller, 2008).

Viele Verfahren der eingeführten Zuordnungsmethoden ziehen die Abbildungsfunktion h zwischen den beiden Modellbeschreibungen m_A und m_B in den Zuordnungsprozess mit ein. Auf diese Weise kann zum Beispiel der Suchraum eingeschränkt werden, was den Berechnungsaufwand entscheidend reduziert. Weiter können mit Hilfe robuster Verfahren Fehlzuzuweisungen detektiert und aus der Abbildungsfunktion eliminiert werden.

2.2.1.3 Algorithmen und Optimierungsansätze

Grundsätzlich kann ein Zuordnungsproblem als Optimierungsaufgabe aufgefasst werden. Nachfolgende Zusammenstellung gibt einen Überblick über bekannte Optimierungsstrategien und Algorithmen.

- **Lokale Optimierung - Gradientenverfahren:**
Die Optimierung erfolgt hier durch ein bekanntes gradientenbasiertes Verfahren. Hierzu gehören die Methode des steilsten Abstiegs, das Newton-Raphson Verfahren sowie das Gauss-Newton Verfahren. Eine Gegenüberstellung der einzelnen Methoden ist in Mautz (2001) zu finden. Die Gauss-Newton Optimierung ist in Anhang B.1 im Detail beschrieben. Voraussetzung für die Anwendbarkeit dieser Verfahren ist die Differenzierbarkeit der zu optimierenden Zielfunktion sowie die Kenntnis genügend genauer Näherungswerte der gesuchten unbekanntenen Parameter.
- **Direkte Optimierungsverfahren:**
Direkte Optimierungsverfahren basieren auf Funktionswertberechnungen der eingeführten Zielfunktion (Mautz, 2001), mit Hilfe derer der Algorithmus entscheidet, welche Teillösungen verworfen, verändert oder gebildet werden. Diese Methode kann für lokale wie auch für globale Optimierungsaufgaben eingesetzt werden. Im Vergleich zu den gradientenbasierten Verfahren zeigen diese Verfahren meist eine geringere Konvergenzgeschwindigkeit. Bekanntes Beispiel dieser Kategorie ist beispielsweise das Downhill-Simplex-Verfahren (Nelder und Mead, 1965).
- **Dynamische Programmierung:**
Die dynamische Programmierung geht von zwei Listen von Objekten aus, für welche ein Pfad zu bestimmen ist, welcher die Objekte aufgrund ihrer Merkmale einander optimal zuordnet. Dieser Pfad bestimmt sich dabei aus der Minimierung der gesamthaft anfallenden Kosten zwischen den Kandidatenpaaren (Baltasvias, 1991). Dieses Verfahren wird vielfach eingesetzt für die eindimensionale Zuordnung von Pixelzeilen in Epipolarbildern für die Generierung von Tiefenbildern.
- **Relaxation:**
Die Merkmale der einander zuzuordnenden Objekte werden bei der Relaxation nach geeigneter Normierung als Wahrscheinlichkeiten interpretiert. Im anschließenden

Zuordnungsprozess werden die Wahrscheinlichkeiten einer Zuordnung schrittweise angepasst und Kandidatenpaare beibehalten, ausgeschlossen oder weiterverarbeitet. Dabei erfolgt die Anpassung der Wahrscheinlichkeiten aufgrund einer Kompatibilitätsfunktion, welche die Eigenschaften benachbarter Merkmale berücksichtigt (Hummel und Zucker, 1987).

- **Informierte- und uninformierte Baumsuche:**
Diese Methode bildet den gesamten Suchraum des Zuordnungsproblems in einen Baum oder Graphen ab. Darin definieren die Kanten zwischen den Knoten des Baumes die anfallenden Kosten bzw. den Nutzen. Durch unterschiedliche Traversierungsstrategien wird derjenige Pfad durch den Suchbaum bestimmt, der das Zuordnungsproblem optimiert (Pearl, 1984). Die informierte Suche nutzt im Gegensatz zur uninformierten Suche zusätzliche Heuristiken, anhand derer der Suchraum eingegrenzt und das Problem effizienter gelöst werden kann.
- **Genetische Algorithmen**
Eine aus der Biologie motivierte Optimierungsstrategie bilden die genetischen Algorithmen, welche die Prinzipien der Evolutionstheorie imitieren. Hierbei wird eine zufällige Zuordnung der Objekte als Individuum bezeichnet und durch ein Chromosom bestehend aus Genen beschrieben. Die Fitness eines Chromosoms entscheidet anschliessend über die Fortpflanzungswahrscheinlichkeit. Durch gezielte Selektion, Mutation und Kombination werden aus den Individuen neue Nachkommen gebildet, welche die starken Gene der Zuordnung erben (Rodehorst, 2004).

2.2.1.4 Zusammenstellung

Die nachfolgend erste Gegenüberstellung (vgl. Tabelle 2-2) zeigt der vielfältige Einsatz der Zuordnungsmethoden für die Lösung von Zuordnungsproblemen auf unterschiedlichen Ebenen. Dabei kann das FBM, wie die Tabelle 2-3 zeigt, durch verschiedenste Algorithmen umgesetzt werden. Interessant ist zudem der Vergleich der eingeführten Zuordnungsebenen mit den in Abschnitt 2.1.2 vorgestellten *Video Registration* Ansätzen. Die eingeführte intensitätsbasierte Georegistrierung lässt sich mit Verfahren der Kategorie Bild-zu-Bild Zuordnung lösen. Demgegenüber greift die 3D-Modellbasierte Georegistrierung auf Verfahren der Kategorie Bild-zu-Modell-Zuordnung zurück.

	ABM	FBM	RM	EFO
Bild-zu-Bild	x	x	x	x
Bild-zu-Modell		x	x	x
Modell-zu-Modell		x	x	

Tabelle 2-2: Gegenüberstellung Zuordnungsebenen vs. Zuordnungsmethoden
(ABM – Flächenbasierte Zuordnung, FBM – Merkmalsbasierte Zuordnung, RM – Relationale Zuordnung, EFO – Energiefunktional-Optimierung)

	ABM	FBM	RM	EFO
Gradientenverfahren	x			(x)
Direkte Optimierung	x	x		x
Dynamische Programmierung		x		(x)
Relaxation		x		
Un- und informierte Suche		x	x	
Genetische Algorithmen		x	x	

Tabelle 2-3: Gegenüberstellung Algorithmen vs. Zuordnungsmethoden

2.2.2 Bild-zu-Modell-Zuordnung und Kanten-Matching

Nachfolgend wird eine Auslese ausgewählter Zuordnungsverfahren und Ansätze behandelt, die für die Lösung der automatischen Bild-zu-3D-Stadtmodell-Zuordnung als Grundlage geeignet erscheinen und auf denen entsprechend aufgebaut werden kann. Die meisten geeigneten Ansätze verwenden hierzu Kantenprimitiven als Merkmale.

2.2.2.1 Merkmalsbasierte Zuordnungsverfahren mit Kantenprimitiven

Ein einfaches kantenbasiertes Verfahren wird von McIntosh und Mutch (1988) vorgeschlagen. McIntosh und Mutch extrahieren zuerst mit einem geeigneten Verfahren Kanten aus den beiden Bildern. Anschliessend werden die Kanten mit Hilfe eines Deskriptors bestehend aus fünf radiometrischen und drei geometrischen Eigenschaften beschrieben. Die radiometrischen Eigenschaften werden aus der Grauwertfunktion in der Umgebung der extrahierten Kante abgeleitet. Als geometrische Eigenschaften werden die Kantenlänge, die Orientierung sowie die Mittelpunktkoordinaten verwendet. Der Zuordnungsprozess berechnet nun aufgrund der Eigenschaften für jedes Merkmal in Bild A einen Ähnlichkeitswert für jedes Merkmal aus Bild B. Für diese Berechnung ist zwischen sämtlichen Eigenschaften eine entsprechende Ähnlichkeitsfunktion definiert, die für die Berechnung des Ähnlichkeitswerts unterschiedlich gewichtet eingeführt werden kann. Kantenpaare werden definitiv einander zugeordnet, wenn diese gegenseitig den höchsten Ähnlichkeitswert aufweisen. Für Details sei auf die originale Publikation von McIntosh und Mutch (1988) verwiesen. Ausblickend sei erwähnt, dass für die Lösung eines Bild-zu-Modell-Zuordnungsproblems lediglich die geometrischen Eigenschaften verwendet werden können. Ein weiterer interessanter Ansatz stammt von Pilu (1997) und löst die Zuordnung von Punktprimitiven in unkalibrierten Stereobildern auf Basis des Algorithmus von Scott und Longuet-Higgins (1991). Dieser Algorithmus verwendet die Singulärwertzerlegung (SVD - engl. *Singular Value Decomposition*) um die korrespondierenden Punkte zu bestimmen.

2.2.2.2 Globale Bild-zu-Modell Zuordnungsverfahren

Einen globalen Ansatz verfolgt Borgefors (1988) zur Lösung von Bild-zu-Modell-Zuordnungsproblemen. Dazu verwendet Borgefors den von Barrow et al. (1977) vorgeschlagenen Chamfer-Algorithmus und erweitert diesen durch ein bildpyramidenbasiertes hierarchisches Zuordnungskonzept und integriert zusätzlich ein entsprechendes Kriterium zum Auffinden von nicht globalen Minima der Zielfunktion. Der Ansatz erlaubt die Zuordnung von Bildaufnahmen auf bekannte in Vektorform vorliegende 3D-Modelle. Dieser Ansatz zählt zu den klassischen Energiefunktional optimierenden Methoden. Der eingeführte Chamfer-Algorithmus ist im Anhang E dargestellt.

Ein ähnliches Verfahren verwenden Kozempel und Reulke (2009) um mit Hilfe einer global definierten Zielfunktion vollautomatisch die äusseren Orientierungsparameter eines Luftbildes aus der automatischen Zuordnung auf ein in Vektorform vorliegendes Strassennetz zu bestimmen. Die Optimierung erfolgt dabei mit Hilfe des Downhill-Simplex-Verfahrens nach Nelder und Mead (1965).

Der von Beveridge und Riseman (1997) unter der Bezeichnung *Local Search* vorgeschlagene Ansatz eignet sich besonders für die Bestimmung der optimalen Zuordnung zweier in 2D definierter Kantensets. Der Ansatz unterstützt zudem Mehrfachzuordnungen. Für die Zuordnung werden zwei Kantensets vorausgesetzt, in welchen jede Kante durch deren Anfangs- und Endpunkt definiert ist. Es wird eine zufällige initiale Zuordnung zwischen den beiden Sets gewählt. Als Abbildungsfunktion zwischen den beiden Kantensets wird eine 2D-Helmerttransformation angenommen. Anschliessend können die Transformationsparameter für diese Abbildung aus der Zuordnung geschätzt und die aktuelle Zuordnung über eine global definierte Zielfunktion beurteilt werden. Die Optimierung dieser Zielfunktion beziehungsweise der Zuordnung erfolgt nun über ein Verfahren des steilsten Abstiegs in der Hemming 1 Nachbarschaft der aktuellen Zuordnung. Für Details über den Algorithmus und das vorgeschlagene Optimierungsverfahren ist auf die Publikation von Beveridge und Riseman (1997) verwiesen.

Ein weiterer möglicher Lösungsansatz bietet die von Ballard (1987) vorgeschlagene Generelle Hough-Transformation (GHT) zur Erkennung von Mustern (PR - engl. *Pattern Recognition*) in Bilddaten. Mit der GHT können in einer Bildaufnahme enthaltene Objekte detektiert und dem entsprechenden Referenzobjekt zugewiesen werden. Sester und Förstner (1989) zeigen eine Anwendung auf, in der mit Hilfe dieser Methode Dachlandschaftspolygone in Luftbildern extrahiert und automatisch erkannt werden. Haala et al. (2002) verwenden die GHT als Basis für die Registrierung von bestehenden 3D-Gebäudemodellen mit Einzelaufnahmen für die Realisierung von *Augmented Reality* Anwendungen. Im Normalfall werden einfache geometrische Grundformen als Referenzobjekte eingesetzt. Grundsätzlich werden bei der GHT die optimalen Abbildungsparameter zwischen den extrahierten Kantenpixeln der Bildaufnahme und des Referenzmusters gesucht. Die Bestimmung der Abbildungsparameter erfolgt ebenfalls über eine global definierte Zielfunktion. Für die Suche des Optimums wird das Zuordnungsproblem in den sogenannten Hough-Raum abgebildet. Der Hough-Raum spannt sich aus der Anzahl unbekannter Parameter der Abbildungs-

funktion auf. Das gesuchte Optimum und dementsprechend die gesuchten Abbildungsparameter können aus dem Hough-Raum extrahiert werden. Das Verfahren ist sehr rechenintensiv und aus diesem Grunde werden vielfach nur Abbildungsfunktionen mit kleinen Freiheitsgraden wie die 2D-Helmerttransformation verwendet.

2.2.2.3 Relationale Zuordnungsverfahren

Dieses Zuordnungsverfahren nutzt im Gegensatz zu den bisher vorgestellten Verfahren zusätzliche aus den Primitiven einer Modellbeschreibung abgeleitete Beziehungen zwischen diesen. Mit Hilfe dieser zusätzlich verfügbaren Information wird es möglich, zwei Modellbeschreibungen aufeinander abzubilden, ohne dass die Abbildungsfunktion, beispielsweise in Form einer relativen Orientierung zwischen zwei Bildern, vorgängig näherungsweise als bekannt vorliegt. Im Zusammenhang mit der relationalen Zuordnung spricht man auch von einem Verfahren, das *low-level Vision* (Bildverarbeitung), *midlevel Vision* (Merkmalsextraktion und Messungen) und *high-level Vision* (Interpretation) kombiniert (Shapiro und Haralick, 1987). Das Verfahren setzt damit voraus, dass mit *high-level* Verfahren aus den Bilddaten entsprechende Modellbeschreibungen abgeleitet werden, die anschliessend interpretiert und einander zugeordnet werden. Das Verfahren eignet sich für die Lösung von Problemstellungen auf unterschiedlichen Zuordnungsebenen (siehe Tabelle 2-2) in denen die Abbildungsfunktion vorgängig nicht bekannt ist. Abbildung 2-3 illustriert das Verfahren anhand einer Bild-zu-Modell-Zuordnung.

Der Ansatz von Shapiro und Haralick

In Shapiro und Haralick (1981) werden Zuordnungsalgorithmen vorgeschlagen, mit welchen zwei Objekte repräsentiert als Modellbeschreibungen bestmöglich aufeinander abgebildet werden. Aufbauend auf dieser Problemformulierung entwickelte sich das heute unter dem Begriff 'Relationale Zuordnung' bekannte Verfahren. Shapiro und Haralick führen zuerst einen Formalismus ein, mit welchem die Modelle und die Abbildungsvorschrift beschrieben werden können. Zudem wird die exakte und nichtexakte Zuordnung unterschieden und definiert. Die nichtexakte Zuordnung unterscheidet sich dahingehend von der exakten, dass nicht davon ausgegangen wird, dass jede Primitive aus der Modellbeschreibung m_A einer Primitiven aus m_B zugeordnet werden kann. Gerade für Anwendungen in der Bildverarbeitung ist die nichtexakte Ausformulierung des Zuordnungsproblems entscheidend. Für die Lösung des nichtexakten Zuordnungsproblems wird eine Baumsuche (engl. *tree search*) vorgeschlagen.

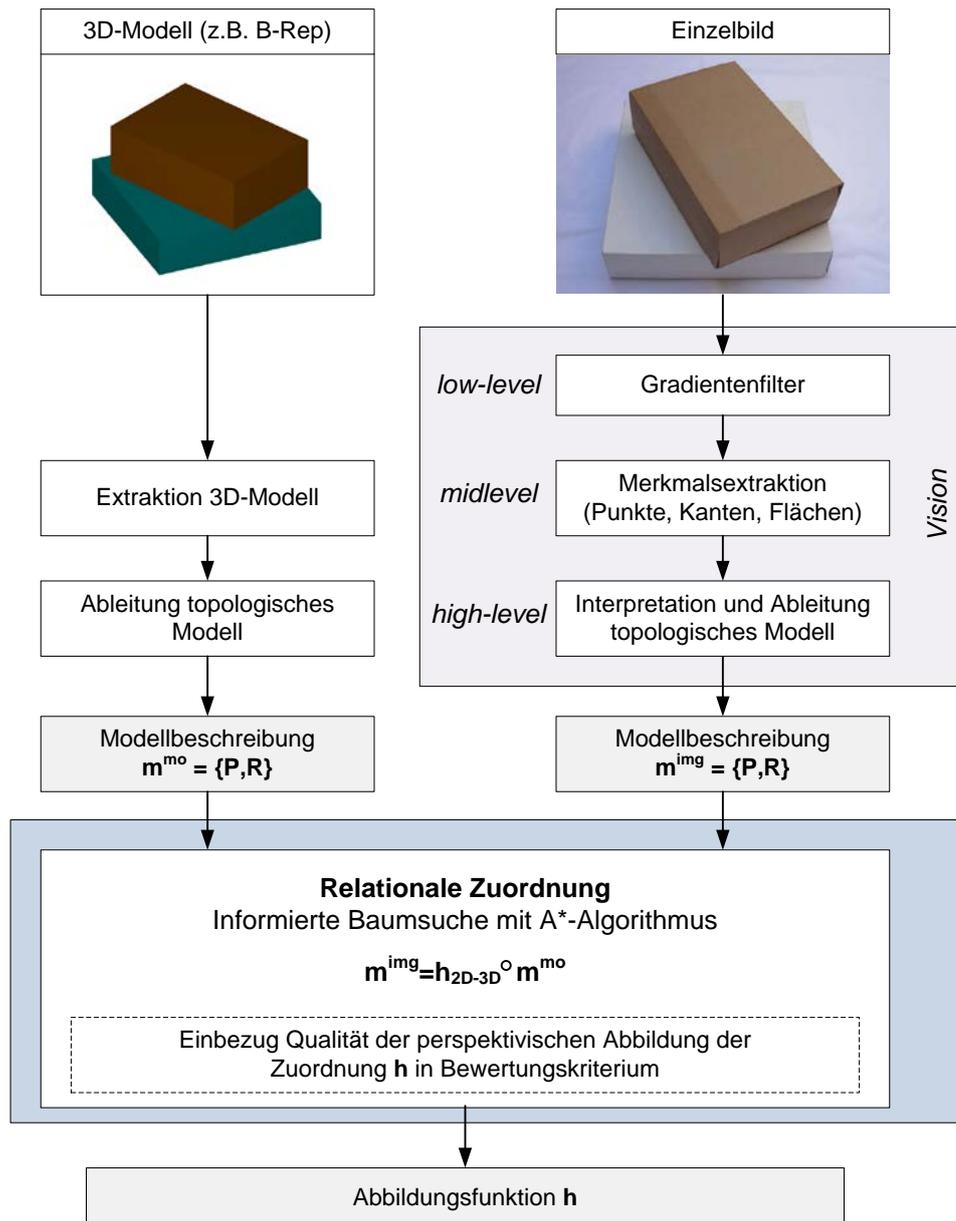


Abbildung 2-3: Schematischer Ablauf relationale Bild-zu-Modell-Zuordnung

Der Ansatz von Vosselman

Vosselman (1992) identifiziert zunächst Defizite des von Shapiro und Haralick vorgeschlagenen Ansatzes, und verbessert das Verfahren, indem die Arbeit das Zuordnungsproblem auf Basis der Informationstheorie formuliert. Das bis anhin verwendete Verfahren zur Beurteilung der Übereinstimmung der einzelnen Primitiven und deren Beziehungen verwendet eine für nichtexakte Zuordnungsprobleme ineffiziente uninformierte Baumsuchmethode (engl. *blind search method*). Zum einen entwirft Vosselman eine auf der Informationstheorie fussende Evaluationsfunktion, mit welcher die Güte einer relationalen Zuordnung zweier Modellbeschreibungen beurteilt werden kann. Als Grundlage verwendet er die von Boyer und Kak (1986) und Boyer und Kak (1988) vorgeschlagene Kostenfunktion, die er zum anderen für das Zuordnungsproblem in eine günstiger anwendbare Nutzenfunktion umformuliert. Für die

Anwendung der Nutzen optimierenden Evaluationsfunktion werden die bekannten Kosten minimierenden Baumsuch-Algorithmen umformuliert. In diesem Zusammenhang wird aufgezeigt, wie die entworfene Evaluationsfunktion in informierten Baumsuch-Algorithmen verwendet wird. Diese Kategorie von Baumsuch-Algorithmen verwendet zusätzlich eine Heuristik die eine effizientere Auswertung erlaubt. Zwei bekannte Algorithmen dieser Kategorie sind die A* Suche (A-Star Suche) oder die ϵ -A* Suche. Das von Vosselman vorgeschlagene RM wird beispielsweise in der Arbeit von Cho (1996) für die automatische äussere Orientierungsbestimmung von Luftbildern eingesetzt.

2.3 Grundlagen zur photogrammetrischen Einzel- und Mehrbildorientierung

In der digitalen photogrammetrischen Produktionskette ist die Bildorientierung eine zentrale Aufgabe, da sie die Voraussetzung bildet für die anschließende Erzeugung diverser Produkte wie digitale Höhenmodelle oder Orthophotos. Zentrale Aufgabe der Bildorientierung ist die Rekonstruktion der Bildaufnahmeposition und -ausrichtung im übergeordneten geodätischen Referenzrahmen. Für Flächenbildsensoren wird dafür die in der Abbildung 2-4 gezeigte perspektivische Abbildung zugrunde gelegt, welche einen Objektpunkt $\mathbf{P} = (X, Y, Z)^T$ im übergeordneten Referenzkoordinatensystem entlang des Aufnahmestrahls in das Bild abbildet \mathbf{p}' . Die Bildorientierungsparameter eines Einzelbildes werden in der Photogrammetrie mit Hilfe des Sensormodells beschrieben, das sich aus der inneren und äusseren Orientierung zusammensetzt. Diese eingeführte Begrifflichkeit wird für diese Arbeit übernommen und wird wie folgt definiert: Die äussere Orientierung beschreibt die räumliche Lage und Orientierung des Bildkoordinatensystems \mathbf{xyz}_{img} bezüglich des übergeordneten Referenzkoordinatensystems \mathbf{XYZ}_{world} mit Hilfe der Parameter:

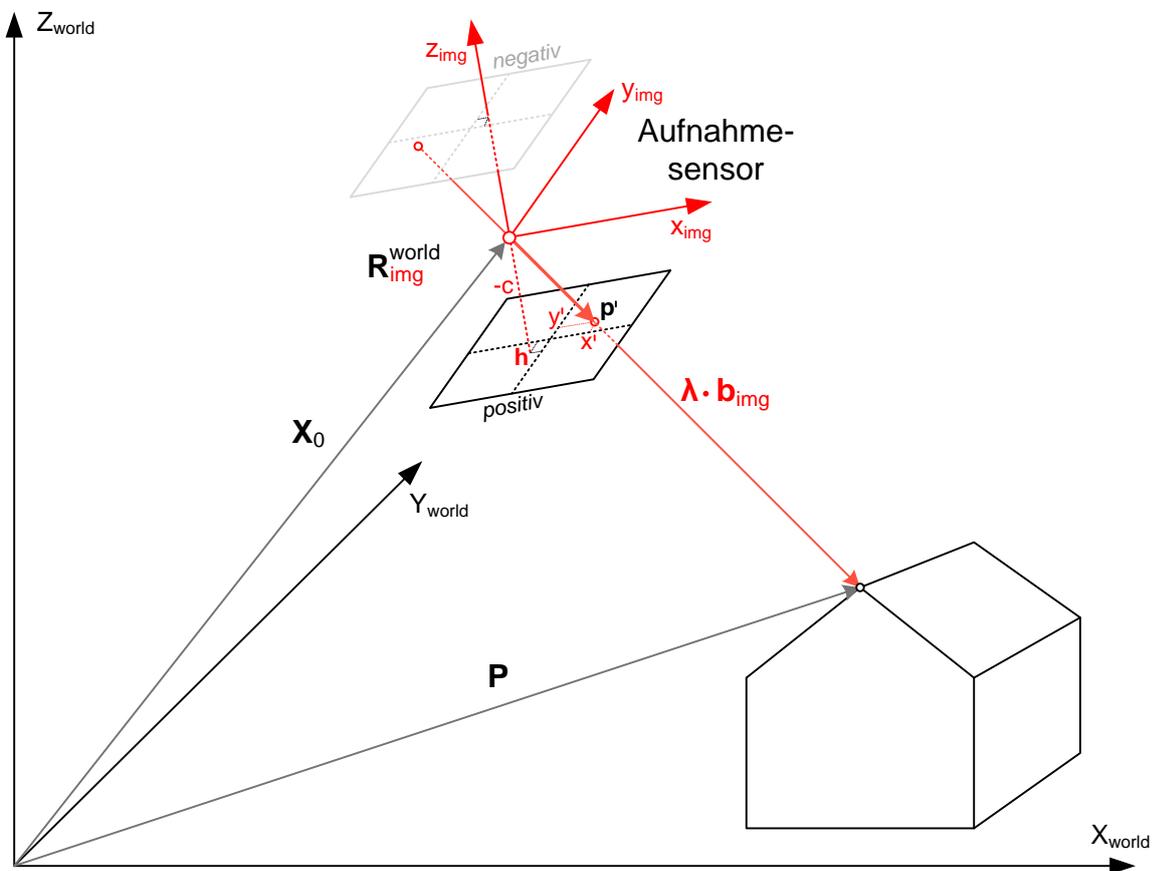


Abbildung 2-4: Perspektivische Transformation und Aufnahmegeometrie

$$\mathbf{X}_0 = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix}$$

Projektionszentrum des Aufnahmesensors im übergeordneten Referenzrahmen.

$$\mathbf{R}_{img}^{world} = R(\omega, \varphi, \kappa)$$

Rotationsmatrix welche die Drehung des Bildkoordinatensystems ins übergeordnete Referenzsystem beschreibt. Die Gesamtdrehung wird aus den drei Eulerwinkeln ω, φ und κ gebildet. (Details zur Parametrisierung der Rotationsmatrix sind in Anhang A zu finden)

Die innere Orientierung definiert das zugrunde gelegte Abbildungsmodell des Aufnahmesensors mit den dazugehörigen Parametern. In dieser Arbeit wird wie in Abbildung 2-4 gezeigt für die verwendeten Flächenbildsensoren die Zentralprojektion zugrunde gelegt, womit die innere Orientierung mit den Parameter

c	Kammerkonstante
x_0, y_0	Bildhauptpunkt \mathbf{h} im Bildkoordinatensystem

festgelegt ist. Für eine digitale Kamera ist im Weiteren vorausgesetzt, dass die Eigenschaften des verwendeten Sensorchips wie Pixelgrösse und Sensordimension bekannt sind. Diese Parameter erlauben die Transformation von Bildpunkten zwischen dem Sensor- und dem Bildkoordinatensystem (Kraus und Waldhäusl, 1997, Kap. 7.2). In den nachfolgenden Betrachtungen wird diese Transformation als bekannt vorausgesetzt und nicht weiter erwähnt. Basierend auf der eingeführten perspektivischen Abbildung lässt sich der ins Bild projizierte Objektpunkt \mathbf{p}' durch folgenden Vektor

$$\mathbf{b}_{img} = \begin{pmatrix} x' - x_0 \\ y' - y_0 \\ -c \end{pmatrix}_{img}$$

im Bildkoordinatensystem durch die Bildkoordinaten x' und y' sowie die Parameter der inneren Orientierung darstellen. In den folgenden Abschnitten werden die in der Photogrammetrie für die Rekonstruktion der Bildorientierung bekannten Verfahren indirekte-, direkte- und integrierte Georeferenzierung rekapituliert und im Anschluss einander gegenübergestellt.

2.3.1 Indirekte Georeferenzierung

Die klassischen aus der Photogrammetrie bekannten Bildorientierungsverfahren basieren auf dem Konzept der indirekten Georeferenzierung. Dazu zählen die aus der einschlägigen Fachliteratur bekannten Methoden Einzelbildorientierung sowie die einstufige und zweistufige Orientierung (Luhmann, 2003, Kap. 4.2). Die Sensormodellrekonstruktion der einzelnen Bildaufnahmen erfolgt bei diesem Ansatz aufgrund

der im übergeordneten geodätischen Referenzrahmen verfügbaren Passpunkte und den in den Aufnahmen durchgeführten Bildmessungen. Die Schätzung der gesuchten äusseren und gegebenenfalls inneren Orientierungsparameter erfolgt bei den Verfahren Einzelbildorientierung und einstufige Orientierung mit einem gradientenbasierten Optimierungsverfahren basierend auf der perspektivischen Abbildung. In diesem Kontext betrachtet ist die Einzelbildorientierung ein Sonderfall der einstufigen Orientierung und entspricht dem räumlichen Rückwärtsschnitt. Im Gegensatz dazu werden im allgemeinen Fall mehrere Bilder gemeinsam orientiert und es wird von einer Bündelblockausgleichung (BA) oder im Luftbildfall von einer Aerotriangulation (AT) gesprochen. Die erwähnte zweistufige Orientierung hingegen orientiert im ersten Schritt die Bilder eines Bildpaares relativ zueinander und transformiert anschliessend im zweiten Schritt die daraus abgeleiteten Modellkoordinaten in den übergeordneten Referenzrahmen. Weiterführende Informationen zu dieser Methode sind in der Fachliteratur beispielsweise in Albertz und Wiggenhagen (2009) zu finden.

Den einstufigen Bildorientierungsverfahren ist die in Abbildung 2-4 dargestellte perspektivische Transformation zugrunde gelegt. Sie kann durch die im Anhang C.1 hergeleiteten Kollinearitätsgleichungen (C-2) ausgedrückt werden. Die Schätzung der unbekannt Parameter erfolgt dabei mit Hilfe der vermittelnden Ausgleichung (vgl. Anhang B). Für Anwendungen in der Photogrammetrie sind die enthaltenen Objektivverzeichnungen in der Aufnahmeoptik bei der Bildorientierung zu berücksichtigen. Bekannte und weit verbreitete Verzeichnungsmodelle nach Brown (1971), Ebner (1976) oder Faig (1975) erlauben die Korrektur dieser Effekte, in dem die Kollinearitätsgleichungen (C-2) um entsprechende Verzeichnungskorrekturterme erweitert werden. Ein weiteres in der Computer Vision verbreitetes Verfahren zur Schätzung und Berücksichtigung der Verzeichnung ist die Methode nach Tsai (1987). Neuere Ansätze verwenden generische Kameramodelle, mit welchen unterschiedliche Kamertypen wie Fischaugenkameras, Katadioptrische Sensoren oder Weitwinkelobjektive modelliert werden können (Luber und Reulke, 2009). Eine detaillierte Abhandlung dieser Thematik ist in Luhmann (2010) zu finden. Diese Thematik ist nicht Bestandteil dieser Arbeit und wird in den weiteren Betrachtungen ausgeklammert. Traditionell beschränkt sich die dargestellte perspektivische Abbildung (C-1) auf die Transformation von Punkten aus dem Objektraum in das Bildkoordinatensystem. In den Veröffentlichungen von Schwermann (1995), Schenk (2004) und Zalmanson (2000) werden Ansätze aufgezeigt, wie die Kollinearitätsgleichungen auf die Abbildung parametrisierter Geometrien und im Speziellen auf Geraden erweitert werden können. Anschliessend werden kurz Spezialitäten der gradientenbasierten Optimierung diskutiert, bevor vertieft auf die beiden Verfahren Einzelbildorientierung und Bündelblockausgleichung eingegangen wird.

2.3.1.1 Parameterschätzung – Methode der kleinsten Quadrate

Für viele Ausgleichsprobleme in der Geodäsie wie auch in der Photogrammetrie wird das gradientenbasierte Optimierungsverfahren nach Gauss-Newton eingesetzt (Mautz, 2001). In der geodätischen Fachliteratur wird dieses Optimierungsverfahren vielfach als

vermittelnde Ausgleichung oder als Ausgleichung nach Gauss-Markov bezeichnet. Dieses Verfahren ist im Anhang B im Detail eingeführt. Gerade für nicht lineare Problemstellungen ist das Gauss-Newton Verfahren sehr effizient und es erreicht in speziellen Fällen quadratische Konvergenz. In Problemstellungen bei denen entweder die zu minimierende Zielfunktion stark gekrümmt ist, ein starkes Messrauschen oder schlechte Näherungswerte vorliegen ist die Konvergenz dieses Verfahrens nicht gegeben. Diese Probleme äussern sich beispielsweise durch eine alternierende oder eine divergierende Iterationsfolge der geschätzten unbekanntes Zuschläge. Eine vertiefte Darstellung der angesprochenen Konvergenzproblematik im Zusammenhang mit geodätischen Ausgleichsproblemen ist in Mautz (2001) diskutiert. Um diesen Unzulänglichkeiten entgegenzuwirken wurde von Levenberg (1944) und Marquardt (1963) das gedämpfte Gauss-Newton Verfahren vorgeschlagen. Diese Methode verwendet im Gegensatz zum klassischen Gauss-Newton Ansatz eine Abstiegs-Schrittweitenregelung, mit welcher die Konvergenz des Iterationsprozesses erzwungen wird. Im optimalen Fall entspricht die Schrittweitenlänge dem Gauss-Newton Verfahren und degradiert in ungünstigen Fällen zum klassischen Gradientenverfahren 'Methode des steilsten Abstiegs' (Madsen et al., 1999, Kap. 2.1). Nachteil der Levenberg-Marquardt Methode ist damit die geringere Konvergenzgeschwindigkeit im Vergleich zum Gauss-Newton Verfahren. Die beiden gradientenbasierten Optimierungsalgorithmen Gauss-Newton und Levenberg-Marquardt sind im Anhang B oder in Press (2007) dargestellt. Weitere in der Geodäsie und Photogrammetrie gebräuchliche und zum Gauss-Markov-Modell verwandte Ausgleichsmodelle sind in Schmid (1977) zusammengestellt.

2.3.1.2 Einzelbildorientierung

Ziel der Einzelbildorientierung ist die Bestimmung der unbekanntes Parameter der äusseren Orientierung $X_0, Y_0, Z_0, \omega, \varphi, \kappa$ eines Einzelbildes aus den Bildmessungen auf bekannte Passpunkte \mathbf{P}_i . Das Problem lässt sich auf Basis der Kollinearitätsgleichungen (C-2) mit einer Ausgleichung nach Gauss-Markov wie in Anhang C.1 oder in Luhmann (2003, Kap. 4.2.3.1.1) aufgezeigt lösen. Da das mathematische Modell bezüglich der unbekanntes Parameter nicht linear ist, setzt das Verfahren entsprechende Näherungswerte für die Linearisierung voraus. Für die Bestimmung werden mindestens drei Bildpunktmessungen auf drei Passpunkte in allgemeiner Lage benötigt. Nach Grafarend et al. (1989) kann das Problem ausgehend von drei Passpunkten in allgemeiner Lage auch mathematisch geschlossen formuliert und ohne die Kenntnis von Näherungswerten streng gelöst werden. Ein weiterer Ansatz ist die von Abdel-Aziz und Karara (1971) vorgeschlagene Direkte Lineare Transformation (DLT), welche für die Bestimmung der gesuchten Parameter keine Näherungswerte voraussetzt. Diese Methode erweitert die Bildkoordinaten um eine Affinitäts-Transformation, womit die Kollinearitätsbeziehung überparametrisiert in linearer Form formuliert werden kann. Die eingeführte Überparametrisierung hat jedoch zur Folge, dass für eine Bestimmung mindestens sechs Passpunkte benötigt werden. Diese

Verfahren eignen sich speziell für die Bestimmung von adäquaten Näherungswerten für die anschließende Linearisierung und Schätzung der Parameter nach Gauss-Newton.

2.3.1.3 Bündelblockausgleichung und Online-Triangulation

Der räumliche Rückwärtsschnitt kann als Spezialfall der Bündelblockausgleichung bezeichnet werden. Im Gegensatz zur Einzelbildorientierung werden hier sämtliche Aufnahmen eines Bildverbandes gemeinsam orientiert. Dafür werden zusätzliche Objektpunkte \mathbf{P}_i in der abgebildeten Szene beobachtet, über welche die Verknüpfung des Bildverbandes erfolgt. Diese Punkte werden in diesem Kontext als Verknüpfungspunkte \mathbf{VP}_i bezeichnet und müssen innerhalb der Bündelblockausgleichung mitgeschätzt werden. Die Bestimmung eines Verknüpfungspunktes erfordert im Minimum die Bildpunktmessung in zwei Bildaufnahmen. Weiter erfordert die Verknüpfung zweier Bilder im Minimum fünf Verknüpfungspunkte. Die Schätzung der unbekannt Parameter erfolgt auf Basis der Kollinearitätsgleichungen (C-2) wie in Anhang C.1 oder Luhmann (2003, Kap. 4.3.2.1) gezeigt nach Gauss-Markov. Wie bei der Einzelbildorientierung ist das mathematische Modell bezüglich der unbekannt Parameter nicht linear und erfordert adäquate Näherungswerte, die mit Hilfe entsprechender Verfahren vorgängig gewonnen werden müssen. Eine Zusammenstellung von Verfahren ist in Luhmann (2003, Kap. 4.3.4) zu finden. Weiter bieten die in Abschnitt 2.1.3 kurz eingeführten Methoden der *Multiple View Geometry* eine voll automatisierbare Alternative zur klassischen photogrammetrischen Näherungswertbestimmung. Der aktuelle Wissensstand über die Bündelblockausgleichung sowie Algorithmen zur Implementierung sind in der Veröffentlichung von Triggs et al. (2000) umfassend zusammengestellt.

Datumsfestlegung

Die reine Verwendung von Verknüpfungspunktbeobachtungen führt in der Ausgleichung zu einem singulären Normalgleichungssystem. Die Form und gegenseitige Lagerung der Aufnahmestrahlenbündel ist zwar bestimmbar, nicht aber die absolute Lage und Orientierung. Dieser Effekt wird als Rang- oder Datumsdefekt bezeichnet und durch die Festlegung der absoluten Lage (3 Translationen), der Orientierung (3 Rotationen) und des Massstabes des Netzes bezüglich des übergeordneten Bezugssystems (Luhmann, 2003, Kap. 4.3.3) behoben. Für die Festlegung werden verschiedene Strategien unterschieden. Im einfachsten Fall werden zusätzliche Bildbeobachtungen auf Passpunkte in die Ausgleichung eingeführt, was auch als harte oder gezwängte Lagerung bezeichnet wird. Für die Festlegung der sieben Datumparameter werden im Minimum zwei Vollpasspunkte sowie ein Höhenpasspunkt benötigt. Weiter besteht die Möglichkeit die Passpunktkoordinaten als zusätzliche Unbekanntbeobachtungen in die Ausgleichung einzuführen, was als weiche- oder auch stochastische Lagerung bezeichnet wird (Kraus et al., 1996, Kap. 4.6.1). Das Resultat der Bündelblockausgleichung mit diesen beiden Varianten ist zum einen beeinflusst durch bestehende Zwänge in den verwendeten Passpunkten und zum andern hängt die aus der Ausgleichung abgeleitete Netzgenauigkeit von der gewählten Passpunktverteilung ab. Die Methode des freien Netzes hingegen lagert die relativ zueinander orientierten

Strahlenbündel auf dem Schwerpunkt der zur Verfügung stehenden Passpunkte. Das Resultat ist damit frei von äusseren Zwängen und eignet sich für die Analyse der erreichten inneren Netzgenauigkeit. Weitere theoretische Grundlagen sowie Implementierungsdetails zur freien Netzlagerung sind in Perelmuter (1979), Wolf (1972) oder Mittermayer (1972) zu finden.

Online-Triangulation und Echtzeit-Verarbeitung

Der Berechnungsaufwand für die Auflösung des Normalgleichungssystems ist speziell bei sehr grossen Bildverbänden hoch und lässt sich durch unterschiedliche Strategien reduzieren. Durch Ausnutzung der vorliegenden Normalgleichungsstruktur kann das Gleichungssystem durch Anwendung des Schurkomplements reduziert und anschliessend mit Hilfe der sogenannten Sparse-Technik oder über die Ausnutzung der Bandstruktur sehr effizient aufgelöst werden (Kraus et al., 1996, Kap. 5.2.2) (Triggs et al., 2000). Untersuchungen wie diejenigen von Engels et al. (2006) haben nachgewiesen, dass mit einer heute üblichen Rechenhardware ein bündelblockbasiertes Echtzeit-Videotracking realisierbar ist. Der untersuchte Algorithmus verwendet das Schurkomplement für eine effiziente Auflösung des Normalgleichungssystems. Beim Einsatz der Bündelblockausgleichung in Echtzeit-Anwendungen wird auch von der Online-Triangulation gesprochen. Die Echtzeitverarbeitung setzt zum einen optimierte Verfahren für die Auflösung des Normalgleichungssystems voraus und benötigt zum anderen sequentielle Algorithmen zur fortlaufenden Integration neuer Beobachtungen in die Bündelblockausgleichung. Grün (1982) verwendet hierfür einen Algorithmus, welcher die Aktualisierung direkt auf dem Gauss faktorisierten Normalgleichungssystem ermöglicht. Weitere Verfahren, welche den Update direkt auf der Gauss/Cholesky-Dreieckszerlegung, den Givens-Transformationen oder auf der Kalman-Kovarianzmatrix durchführen, sind in (Grün, 1985) zusammengestellt. Wie bereits in Abschnitt 2.1.4 aufgezeigt, beschäftigt sich die Robotik bei der Umsetzung von SLAM-Anwendungen mit sehr ähnlichen Fragestellungen. Im Gegensatz zu den Online-Triangulationsansätzen werden dort vielfach KF-basierte sequentielle Zustandsschätzer propagiert.

2.3.2 Direkte Georeferenzierung

Ziel der direkten Georeferenzierung ist das direkte Messen der gesuchten äusseren Orientierung eines beliebigen Aufnahmesensors (z.B. Flächenbildsensor, Linienscanner oder LiDAR) montiert auf einer mobilen Plattform mit Hilfe eines *navigation grade* INS in Kombination mit einem geodätischen GNSS-Empfänger. Die dafür benötigte KF-basierte Datenfusion sowie Details zu den verwendeten Sensoren werden in den Abschnitten 2.4 und 2.5.3 besprochen. Damit erlaubt dieser Ansatz die direkte Bestimmung gesuchter Objektpunktkoordinaten \mathbf{P} im übergeordneten geodätischen Referenzsystem. Abbildung 2-5 zeigt die Aufnahmegeometrie im übergeordneten geodätischen Bezugssystem. Die Grundgleichung der direkten Georeferenzierung (C-5) dargestellt im Anhang C.2 erlaubt die Bestimmung der gesuchten 3D-Objektpunktkoordinaten aus den direkt gemessenen äusseren Orientierungen (Cramer, 2001a). Aus der INS/GNSS-basierten Sensordatenfusion resultieren die Terme \mathbf{X}_{world} und $\mathbf{R}_{body}^{world}$. Die erreichbare absolute Genauigkeit der auf diese Weise bestimmten Aufnahme-trajektorie wird im Wesentlichen durch die Qualität der GNSS-Lösung bestimmt.

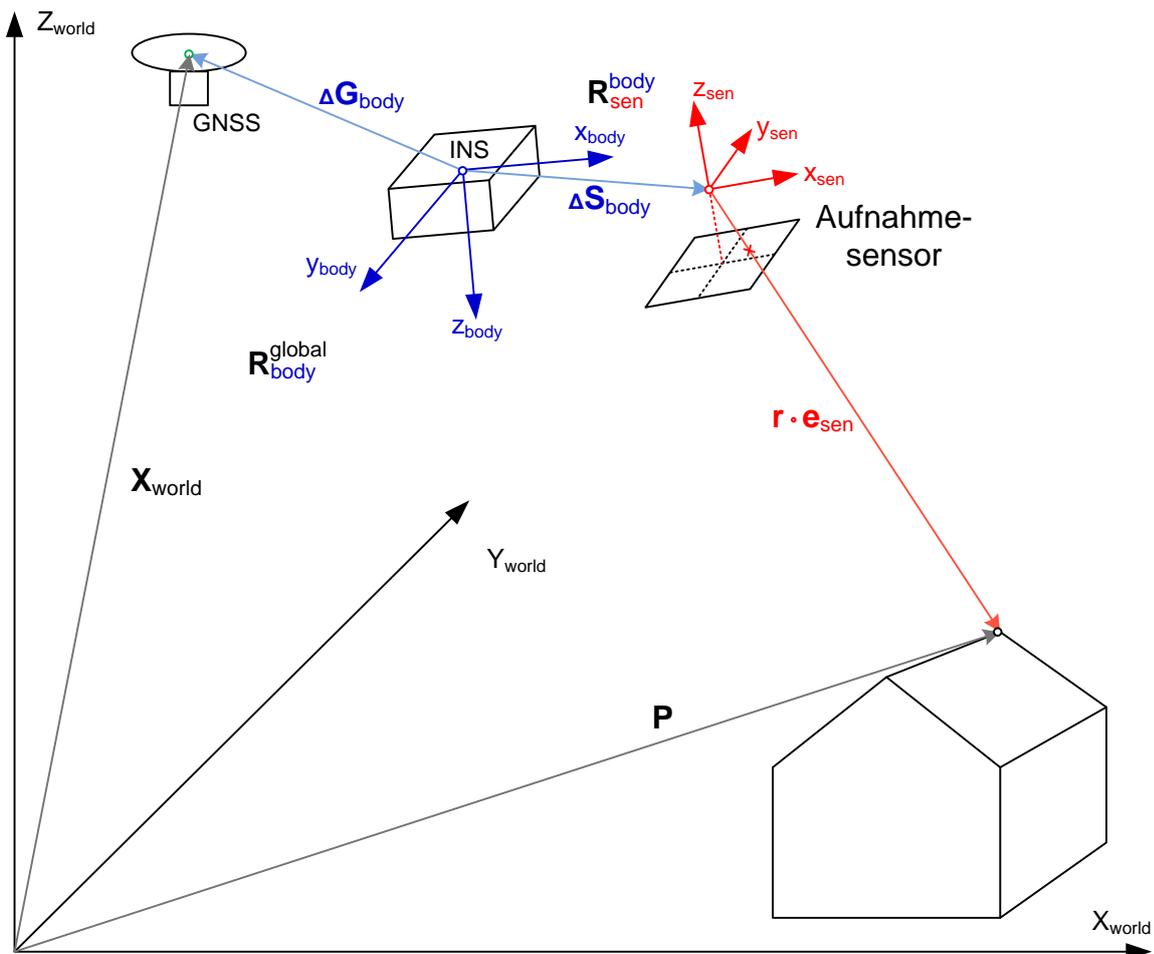


Abbildung 2-5: Aufnahmegeometrie der direkten Georeferenzierung

Die relative Genauigkeit hängt hingegen hauptsächlich von der Qualität des eingesetzten INS ab. Der verwendete Aufnahmesensor liefert zusätzlich die Terme r und \mathbf{e}_{sen} . Bei der Verwendung eines Video- bzw. eines Flächenbildsensors wird davon ausgegangen, dass die Parameter der inneren Orientierung in Form eines korrekten Kameramodells (z.B. Kameramodell nach Brown (1971)) bestimmt wurden. Die weiteren für die Bestimmung notwendigen Terme \mathbf{R}_{sen}^{body} , $\Delta\mathbf{S}_{body}$ und $\Delta\mathbf{G}_{body}$ müssen mit Hilfe einer vorgängig durchgeführten Systemkalibrierung bestimmt und festgelegt werden.

2.3.2.1 Systemkalibrierung

Mit Hilfe der Aufnahmesystemkalibrierung wird die Beziehung zwischen dem Aufnahmesensorsystem, dem INS-Körpersystem (Plattformkoordinatensystem) und dem GNSS-Antennenphasenzentrum festgelegt. Die Fehlausrichtung (engl. *boresight misalignment*) \mathbf{R}_{sen}^{body} definiert die Verdrehung zwischen Aufnahmesensorkoordinatensystem und INS-Körpersystem. Die beiden Translationen $\Delta\mathbf{S}_{body}$ und $\Delta\mathbf{G}_{body}$ werden als Hebelarmkomponenten (engl. *lever-arm*) bezeichnet und beschreiben die räumliche Lage des Aufnahmesensorsystem-Referenzpunktes und des GNSS-Antennenphasenzentrums bezüglich des INS-Referenzpunktes. Nachfolgend sind kurz die Verfahren zur Bestimmung dieser Komponenten rekapituliert.

Fehlausrichtung

Die Fehlausrichtung kann nicht direkt gemessen werden und wird meist aus dem Vergleich zwischen direkt und indirekt bestimmter Orientierungslösung abgeleitet. Für diesen Zweck wird bei luftgestützten Plattformen ein sogenannter Kalibrierflug durchgeführt. Damit die gesuchten Drehwinkel zuverlässig geschätzt werden können, empfiehlt es sich ein Kreuzflug über einem bekannten Passpunktfeld durchzuführen. Aufgrund dieses Flugmusters lässt sich bei der späteren Auswertung die Fehlausrichtung von der äusseren Orientierung bestmöglich dekorrelieren. Die Schätzung der Drehwinkel erfolgt meist durch eine der folgenden beiden Methoden.

- Drehwinkel aus Rotationsmatrixdifferenz:
Bei dieser Methode werden die Bildorientierungen der Einzelbilder des Kreuzfluges mittels direkter und indirekter Georeferenzierung unabhängig bestimmt. Die Differenz zwischen den beiden unabhängig bestimmten Orientierungen bildet die gesuchte Fehlausrichtung. Dabei muss beachtet werden, dass die Differenzbildung und Mittelbildung in der Eulerwinkelrepräsentation erfolgt. Weiterführende Informationen zu diesem Ansatz sind in Skaloud (1999b) und Bäumker (2007) zu finden.
- Mitschätzung der Drehwinkel in AT:
Die direkt gemessenen äusseren Orientierungen der Kreuzflugeinzelbilder können bei diesem Ansatz als zusätzliche Beobachtungen in eine Aerotriangulation eingeführt werden. Für die Schätzung müssen jedoch die Kollinearitätsgleichungen mit den drei unbekanntem Fehlausrichtungswinkeln erweitert werden. Anschliessend können diese im Gesamtausgleich mitgeschätzt werden (Stallmann, 2008).

Hebelarmkomponenten

Die unbekanntenen Translationen können meist an der Plattform selbst mit Hilfe geodätischer Messinstrumente direkt eingemessen werden. Vielfach kann die Verschiebung $\Delta \mathbf{G}_{body}$ zwischen GNSS-Antennenphasenzentrum und INS-Körpersystem bereits in der Navigationsrechnung aus INS und GNSS berücksichtigt werden, womit dieser Term in Gleichung (C-5) entfällt.

2.3.2.2 Probleme und systematische Fehlereinflüsse

Die erreichbare Objektpunktgenauigkeit wird bei der direkten Georeferenzierung durch verschiedene Faktoren beeinflusst. Optimale Resultate können erreicht werden, wenn folgende Probleme und Einflüsse bei der Konzeption des luftgestützten Aufnahmesystems und der Auswertung berücksichtigt werden (Skaloud, 1999b).

- **Sensoranordnung auf dem Aufnahmesystem:**
Die GNSS-Antenne, das INS und der Aufnahmesensor sollten wenn möglich mit geringem Abstand zueinander auf der Aufnahmeplattform angeordnet werden, damit der Hebelarm-Einfluss möglichst klein gehalten werden kann. Die Sensoren müssen möglichst rigide zueinander montiert werden, damit Fehlausrichtung und Hebelarmkomponenten konstant bleiben. Bessere Resultate werden erreicht, wenn der Aufnahmesensor auf einer kreiselstabilisierten Plattform eingesetzt wird, damit Vibrationen der Trägerplattform möglichst entkoppelt werden. In diesem Fall muss jedoch das INS direkt am Aufnahmesensor angebracht und die GNSS-Antenne möglichst senkrecht darüber installiert werden.
- **Sensor-Synchronisation:**
Die einzelnen Sensoren müssen genau bezüglich derselben Zeitreferenz betrieben werden. Die benötigte Synchronisationsgenauigkeit hängt von der Aufnahmeplattformdynamik ab. Vielfach wird die koordinierte Weltzeit (UTC - engl. *Universal Time Coordinated*) als Zeitreferenz verwendet, welche mit hoher Genauigkeit durch GNSS realisiert werden kann. Bei der Synchronisation (Triggerung - Bildauslösung) der einzelnen Sensoren müssen gegebenenfalls Hardware bedingte Signalverzögerungen berücksichtigt werden.
- **Initialisierung Inertialnavigationssystem:**
Die Initialisierung wird benötigt um das INS-Körpersystem bezüglich des übergeordneten geodätischen Referenzsystems für die anschließende Navigation zu orientieren. In der Fachliteratur werden die Verfahren statische und dynamische Initialisierung unterschieden. Für luftgestützte Photogrammetrieanwendungen wird das dynamische Verfahren verwendet. Das dynamische Verfahren kann im Gegensatz zum statischen auch für INS geringer Qualität verwendet werden. Das dynamische Verfahren macht sich die durch GNSS ebenfalls beobachtbare Systemdynamik für die Initialisierung zu Nutze. Zu erwähnen bleibt: in einem INS/GNSS integrierten Navigationssystem wird die Initialisierung streng nie abgeschlossen und fortlaufend im Systemzustand mitgeschätzt.

2.3.3 Integrierte Georeferenzierung

Im Unterschied zum Ansatz der direkten Georeferenzierung werden bei der integrierten oder kombinierten Georeferenzierung zu den direkt aus IMU und GNSS bestimmten äusseren Orientierungselementen zusätzliche Bildbeobachtungen auf bekannte Pass- und Verknüpfungspunkte zur bestmöglichen Bestimmung dieser Parameter mitverwendet. Die zusätzliche Integration des indirekten Georeferenzierungsansatzes erlaubt die Schätzung und Berücksichtigung verbleibender systematischer Fehler der direkt gemessenen äusseren Orientierungsparameter. Die direkt gemessene Flugtrajektorie ist unter anderem mit unterschiedlichen systematischen Fehlern, beispielsweise verursacht durch falsch fixierte Phasenmehrdeutigkeiten nach Signalunterbrüchen (*Cycle Slips*) in der kinematischen GNSS-Lösung, überlagert (Haala, 2005) (Favey, 2001) (Cocard, 1995). Für die Schätzung und Korrektur dieser Systematiken in der Luftbildphotogrammetrie diskutieren Haala (2005) und Cramer (2001a) die Verfahren 'streifenweiser Integrationsansatz' und 'Orientierungsbildansatz'. Der streifenweise Integrationsansatz schätzt aus den zusätzlich verfügbaren Bildbeobachtungen für die direkt gemessenen Positions- und Orientierungsdaten einen Shift und einen zeitabhängigen Drift. Die Shift- und Driftparameter werden dabei für einen Flugstreifen einzeln oder gemeinsam für den gesamten Bildblock geschätzt. Alternativ dazu unterteilt der Orientierungsbildansatz die Flugstreifen in noch viel kleinere Abschnitte, für welche jeweils ein eigener Korrekturparametersatz definiert und geschätzt wird. Weitere integrierte Datenfusionsansätze für die gemeinsame Verarbeitung von GNSS und Bildbeobachtungen werden in Ellum und El-Sheimy (2006) vorgestellt und verglichen. Implizit werden hier die beiden Integrationskonzepte 'INS/GNSS/Vision-basierte Schätzung der Aufnahmetrajektorie' und 'INS/GNSS-unterstützte AT' unterschieden. Im Gegensatz zu den vorgeschlagenen Verfahren von Haala und Cramer für die INS/GNSS/Vision-basierte Schätzung der Aufnahmetrajektorie verwenden Ellum und El-Sheimy dazu einen Kalman-Filter. Diese Integrationsstrategie steht in engem Zusammenhang mit den Konzepten, die in der Robotik für die Lösung des SLAM-Problems genutzt werden. Im Gegensatz zum ersten Integrationskonzept werden bei der INS/GNSS-unterstützten AT die direkt gemessenen äusseren Orientierungen als zusätzliche Beobachtungen in die Bündelblockausgleichung eingeführt. Diese Integrationsstrategie ermöglicht eine erhebliche Reduktion der notwendigen Lage- und Höhenpasspunkte im Bildblock für das Erreichen einer vergleichbaren Georeferenzierungsgenauigkeit im Vergleich zu einer klassischen Aerotriangulation. Zu berücksichtigen gilt, dass dieses Verfahren jedoch auf Bildflüge mit Flächenbildsensoren beschränkt ist. Viele Autoren wie beispielsweise Heipke et al. (2002) unterscheiden die beiden integrierten Georeferenzierungskonzepte nicht streng.

2.3.4 Gegenüberstellung der Bildorientierungsansätze

Die drei in diesem Kapitel diskutierten Ansätze zur Rekonstruktion der Aufnahmebildorientierung in der Photogrammetrie sind in Abbildung 2-6 einander gegenübergestellt. Oben in der Abbildung sind die einzelnen für die Bildorientierung benötigten Grundlagedaten aufgeführt. Für den Vergleich der Ansätze ist folgende noch nicht eingeführte Betrachtungsweise interessant. In der indirekten Georeferenzierung werden die gesuchten Orientierungsparameter sowie die zusätzlich zu bestimmenden Verknüpfungspunkte, mit Hilfe der aus den Bildbeobachtungen abgeleiteten Strahlenbündel, in die Passpunkte im übergeordneten Referenzrahmen interpoliert. Die direkte Georeferenzierung hingegen kann in diesem Kontext als extrapolierte Lösung betrachtet werden.

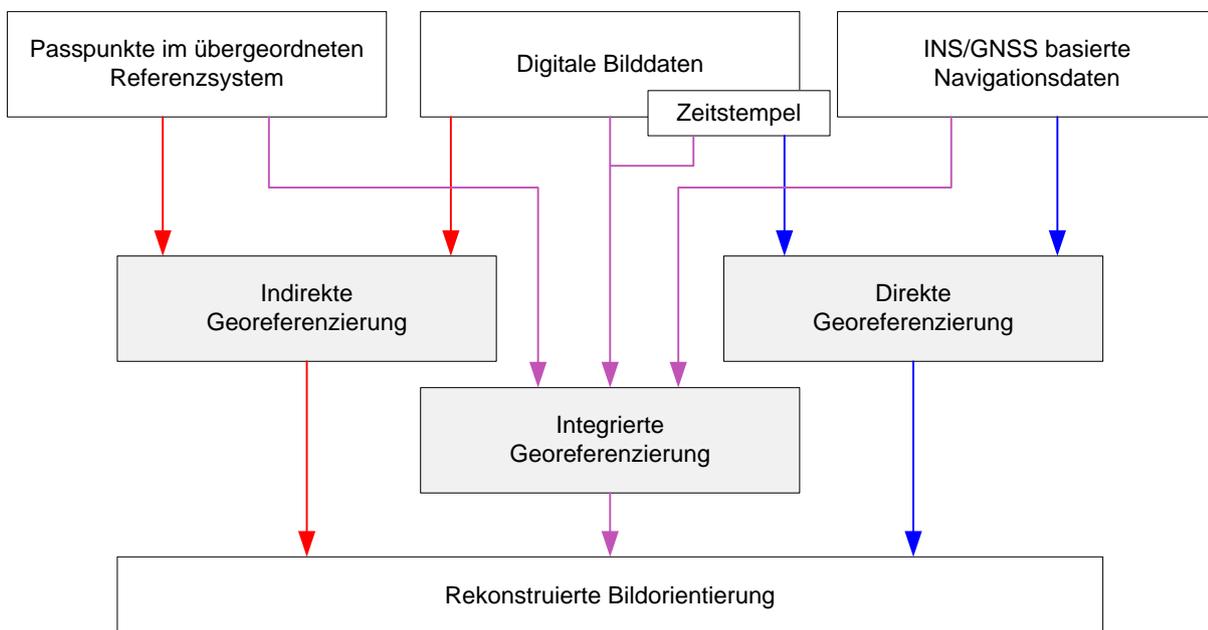


Abbildung 2-6: Gegenüberstellung photogrammetrische Bildorientierungsansätze

Die Bildorientierungen werden hier bezüglich einer im übergeordneten Referenzrahmen vorliegenden GNSS-Referenzstation absolut bestimmt, wodurch sich sämtliche systematischen Einflüsse der INS/GNSS-Messtechnik auf die resultierende Lösung übertragen. Bei der integrierten Lösung wird nun, vereinfacht ausgedrückt, die durch Extrapolation bestimmte Aufnahmetrajektorie in zusätzliche im übergeordneten Referenzrahmen zur Verfügung stehende Passpunkte transformiert. Dadurch können die systematischen Einflüsse der Extrapolation bestmöglich kompensiert werden. Die direkte Georeferenzierung ermöglicht im Vergleich zur indirekten Lösung, aufgrund der vom INS hochfrequent bestimmten Navigationsdaten, die Orientierung von zeilenbasierten Bildsensoren oder LiDAR-Daten. Eine abschliessende Gegenüberstellung der drei Orientierungsansätze hinsichtlich erreichbarer Objektpunktgenauigkeiten in der Luftbildphotogrammetrie sind in Cramer (2001b) zu finden.

2.4 Grundlagen zur Kalman-Filter-basierten Sensordatenfusion

Ziel der Kalman-Filter-basierten Sensordatenfusion ist die kontinuierliche optimale Schätzung des gesuchten Systemzustandes (engl. *state vector*), beispielsweise die Fluglage einer Drohne, aus Beobachtungen unterschiedlicher Sensoren. Der von Kalman (1960) eingeführte lineare zeitdiskrete rekursive Algorithmus, schätzt den Systemzustand \mathbf{x}_t sowie die dazugehörige Fehlerkovarianzmatrix \mathbf{P}_t zum Zeitpunkt t aufgrund der zu den Epochen $1, \dots, t$ zur Verfügung stehenden Beobachtungen $\mathbf{z}_1, \dots, \mathbf{z}_t$ optimal im Sinne der minimalen Schätzfehlervarianz. Das Verhalten des Systemzustandes wird im Kalman-Filter anhand eines Zufallsprozesses in zeitdiskreter Form wie folgt beschrieben:

$$\mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{w}_t \quad (8)$$

Darin definiert $\mathbf{\Phi}_t$ die Transitionsmatrix für den Übergang des Systemzustandes \mathbf{x} von der Epoche t zur Epoche $t+1$. Der Term \mathbf{w}_t beschreibt dabei das System- bzw. das Prozessrauschen. Die Beobachtung dieses Prozesses erfolgt mit

$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t \quad (9)$$

basierend auf den zu diskreten Zeitpunkten zur Verfügung stehenden Messungen \mathbf{z}_t . Die Designmatrix \mathbf{H}_t bildet die Beobachtungen des Zeitpunktes t auf den Systemzustand \mathbf{x}_t ab und \mathbf{v}_t beschreibt das Messrauschen der Beobachtungen. Das eingeführte Prozess- und Messrauschen wird als nicht-kreuzkorreliertes mittelwertfreies weisses Rauschen mit den Kovarianzmatrizen \mathbf{Q}_t für das Prozess- und \mathbf{R}_t für das Messrauschen angenommen. Im Weiteren ist vorausgesetzt, dass die Systemdynamik und die entsprechenden Beobachtungsgleichungen bezüglich des Systemzustandes linear sind. Ist dies nicht gegeben, muss ein alternatives Filtermodell, wie beispielsweise das in Brown und Hwang (1997, Kap. 9) beschriebene Erweiterte Kalman-Filter (EKF) oder das von Julier und Uhlmann (1997) eingeführte *Unscented* Kalman-Filter (UKF), für die Lösung verwendet werden. Aus den eingeführten Zustandsdarstellungsgleichungen folgt der in Abbildung 2-7 in abstrakter Form dargestellte Kalman-Filter-Algorithmus. Das Kalman-Filter wird mit einem angenommenem Systemzustand und dessen Fehlervarianz initialisiert. Die damit a priori vorliegende Systemzustandsbeschreibung $\mathbf{x}_t^{(-)}$ und $\mathbf{P}_t^{(-)}$ wird nun mit den zum Zeitpunkt t zur Verfügung stehenden Beobachtungen \mathbf{z}_t aktualisiert. Dieser Schritt wird auch als Korrektur (engl. *update*) bezeichnet und liefert die a posteriori optimal geschätzte Systemzustandsbeschreibung $\mathbf{x}_t^{(+)}$ und $\mathbf{P}_t^{(+)}$.

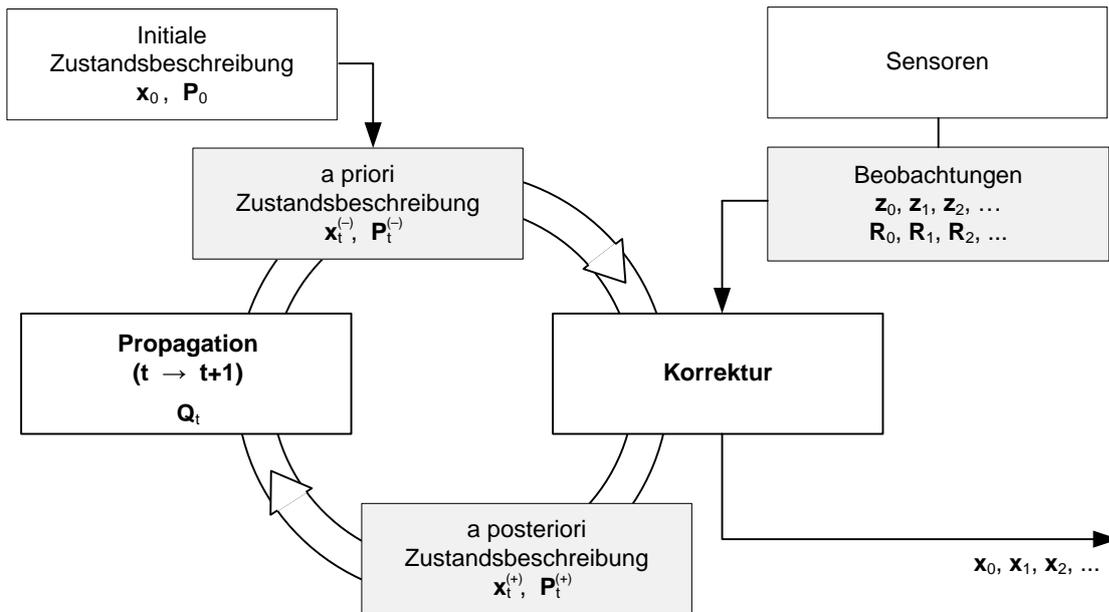


Abbildung 2-7: Prinzip zeitdiskretes Kalman-Filter

Die anschließende Propagation (engl. *prediction*) prädiziert den a posteriori vorliegenden Systemzustand aufgrund der verwendeten Systemdynamik auf die Zustandsbeschreibung $\mathbf{x}_{t+1}^{(-)}$ und $\mathbf{P}_{t+1}^{(-)}$ der Epoche $t+1$. Eine Einführung in das beschriebene Kalman-Filter ist beispielsweise in Welch und Bishop (2004) zu finden. Für weiterführende Informationen zur Herleitung des zeitdiskreten Kalman-Filters ist auf die einschlägige Fachliteratur (Brown und Hwang, 1997) oder (Gelb, 1986) verwiesen. Eine mathematische Ausformulierung des eingeführten diskreten Kalman-Filter-Algorithmus ist in Anhang D zu finden.

In der Literatur werden unterschiedliche Filterarchitekturen und Ansätze unterschieden. Aufgrund der Problemstellung und der zur Verfügung stehenden Sensoren können das *full state* Filter und das *error state* Filter unterschieden werden (Winkler, 2007). In der *full state* Architektur enthält der Zustandsvektor des Filters direkt die zu bestimmenden gesuchten unbekannt Parameter aus der zur Verfügung stehenden Sensorkonfiguration. Im Gegensatz dazu verwendet das *error state* Filter meist einen Primärsensor, aus welchem die gesuchten unbekannt Parameter direkt abgeleitet werden. Der Zustandsvektor besteht in diesem Fall aus den aktuellen systematischen Fehlern dieses Sensors, welche anschliessend im Primärsensor berücksichtigt werden. Die Schätzung dieses Zustandsvektors erfolgt dabei aufgrund der zusätzlich verfügbaren Beobachtungen der weiteren Sensoren. Bei der Berücksichtigung des geschätzten Fehlerzustandes wird zwischen den Konzepten *feedforward* und *feedback* unterschieden (Cramer, 2001a). Beim *feedforward* Design wird der geschätzte Fehler direkt an den vom Primärsensor abgeleiteten gesuchten Parametern berücksichtigt. Demgegenüber führt das *feedback* Filter den aktuellen Fehlerzustand direkt in den Primärsensor zurück, aufgrund dessen die Rohmesswerte korrigiert und verbesserte abgeleitete Parameter resultieren. In diesem Zusammenhang steht auch die Unterscheidung einer *tightly coupled* und *loosely coupled* Datenintegration. Das *loosely coupled* Filter verwendet für die Zustandsschätzung bereits von den einzelnen

Sensoren vorverarbeitete Beobachtungen. Im Gegensatz dazu verarbeitet das *tightly coupled* Filter direkt die verfügbaren Rohmesswerte und setzt diese in Beziehung zu den unbekanntem Parametern im Zustandsvektor, wodurch eine optimale Güte der Zustandsschätzung erzielt wird. Demgegenüber lässt sich jedoch das *loosely coupled* Filter sehr flexibel einsetzen, da die Systemarchitektur sehr einfach durch zusätzliche Sensoren ergänzt oder einzelne Sensoren aus der Architektur entfernt werden können.

2.5 Basistechnologien

Dieses Kapitel gibt eine Einführung in relevante Technologien, welche die Umsetzung dieser Arbeit unterstützen. Der erste Abschnitt vermittelt einen kurzen Überblick über virtuelle Globen und zeigt wie diese Technologie für die effiziente Beantwortung räumlicher Anfragen auf bestehenden Geodatenbeständen genutzt werden kann. Nachfolgender Abschnitt rekapituliert unbemannte Flugsysteme und diskutiert den Aspekt der Fluglagebestimmung und -regelung. Aufbauend darauf behandelt der Abschnitt 'Integrierte Navigationssysteme' unterschiedliche verwendete Navigationssensoren welche bei der Fluglagebestimmung auf Mini- und Mikrodrohnen sowie bei der hochpräzisen Bestimmung der Aufnahmetrajektorie in der Luftbildphotogrammetrie Verwendung finden. Der abschliessende Abschnitt widmet sich der Videotechnologie und der Datenübertragung, welche speziell im Zusammenhang mit der Echtzeit-Videodatenintegration steht.

2.5.1 Virtuelle Globen

In den letzten Jahren haben sich virtuelle Globen und damit realisierte 3D-Geoinformationsdienste auf dem Massenmarkt etabliert, bekannte Beispiele dafür sind die beiden Anwendungen Google Earth und Microsoft Bing Maps 3D. Die intensive Entwicklung dieser Technologien hat bereits vor gut einer Dekade eingesetzt und interaktive Terrainvisualisierungstechnologien (engl. *Terrain Engines*) haben sich über diese Periode von einfachen Flache-Erde unterstützenden zu komplexen auf geodätischen Referenzsystemen basierenden Lösungen entwickelt. Mit dem Schritt von der flachen Erde auf den Globus oder in geodätischem Sinne aus der Kartenprojektion auf ein globales Referenzellipsoid hat sich auch der Begriff virtueller Globus oder synonym dafür digitaler Globus (engl. *Digital Globe*) für diese Art von *Terrain Engine* etabliert. Die heutigen virtuellen Globen beeindrucken mit ihrer Fähigkeit beliebig grosse Geodatenätze in Echtzeit visualisieren zu können. Üblicherweise werden Orthophotomosaiken, digitale Höhenmodelle, 3D-Statdtmodelle sowie Datensätze mit interessanten Orten (POI - engl. *Point of Interest*) gesamter Länder oder Kontinente unterstützt. Diese riesigen Datenmengen erfordern für eine Visualisierung in Echtzeit ein kontinuierliches Nachladen der entsprechenden Geodaten in die *Terrain Engine* mit Hilfe spezieller auf räumlichen und mehrfach aufgelösten Datenstrukturen basierten

Streaming-Ansätzen. Einen Überblick über den aktuellen Stand der Technologie und Nutzung dieser Dienste ist in Nebiker et al. (2010) zu finden. Speziell im Hinblick auf die gesteckten Ziele dieser Arbeit sind die meisten virtuellen Globen nicht dafür ausgelegt, beliebige in Echtzeit erfasste Geodaten zu integrieren und dem Nutzer interaktiv zur Verfügung zu stellen. Bei vielen Technologien sind die zur Verfügung stehende Funktionen und Programmierschnittstellen bezüglich der Echtzeitintegration von Geodaten sehr limitiert. Vielfach sind die zugrunde gelegten geodätischen Referenzsysteme nicht publiziert, womit sich eine automatisierte Echtzeitintegration von Geodaten mit ausreichender Georeferenzierungsgenauigkeit ohne manuelle Nachbearbeitung problematisch gestaltet. Aus diesen Gründen wird für diese Arbeit auf den sehr flexibel einsetzbaren an der Fachhochschule Nordwestschweiz entwickelten virtuellen Globus i3D zurückgegriffen (Christen und Nebiker, 2010). Die Verwendung der i3D-Technologie im Rahmen dieser Arbeit ermöglicht im Weiteren eine neue Nutzungsmöglichkeit virtueller Globen aufzuzeigen. Basierend auf den implementierten *Streaming*-Ansätzen in Kombination mit Algorithmen aus der Computergraphik eignet sich diese Technologie für die sehr effiziente Beantwortung räumlicher *View Frustum* Anfragen kombiniert mit der zusätzlich verfügbaren Semantik auf einem beliebig grossen Geodatenbestand. Der Einsatz als *3D Spatial Query Engine* stellt in Bezug auf die Nutzung virtueller Globen ein neues Anwendungsparadigma dar, weg von der reinen Präsentation von Geodaten, hin zur intelligenten digitalen Abbildung unserer realen Welt.

2.5.2 Unbemannte Flugsysteme

Für unbemannte Flugsysteme wird vielfach der englische Ausdruck *Unmanned Aircraft System* (UAS) verwendet. Unter der Bezeichnung UAS definiert das *U.S. Department of Defense* und das *U.S. Department of Transportation* eine vollständige Systemlösung, die den Betrieb eines unbemannten Fluggeräts gewährleistet (Weatherington, 2005). Ein UAS besteht typischerweise aus der unbemannten Flugplattform mit Navigationseinheit, den zusätzlichen Nutzsensoren z.B. zur Erfassung von Geodaten, der Bodenkontrollstation und aus Telemetrie- und Datenübertragungskomponenten. Die unbemannte Flugplattform oder das Fluggerät selbst wird in dieser Terminologie als *Unmanned Aircraft* (UA) bezeichnet, welche mit Hilfe der Bodenkontrollstation vollständig autonom oder ferngesteuert geflogen werden kann. Heute existiert gerade auch auf dem zivilen Markt eine grosse Anzahl unbemannter Systeme in unterschiedlichen Kategorien. In dieser Arbeit kommen Mini- und Mikroflugplattformen zur Anwendung. Das sind Systeme die ein maximales Abfluggewicht von 30kg bzw. 5kg nicht überschreiten. Die von *UVS International*¹ aktuell gültige Klassifizierung der unterschiedlichen Flugplattformen ist in Bento (2008) zusammengestellt. In Bezug auf drohnenbasierte photogrammetrische Anwendungen führt Eisenbeiss (2009) eine angepasste Klassifizierung der Systeme ein. Der Kern der autonomen unbemannten

¹ Unmanned Vehicle Systems International (<http://www.uvs-international.org>)

Flugplattform bildet die Navigationseinheit, die vielfach auch als Autopilot bezeichnet wird. Die Bestimmung des aktuellen Fluglagezustands ist eine zentrale Aufgabe des Fluglagekontrollers innerhalb der Navigationseinheit. Aufgrund des aktuell geschätzten Fluglagezustandes können automatisch entsprechende plattformabhängige mechanische Steuerbefehle bestimmt werden, die es erlauben die gewünschte Fluglage automatisch beizubehalten. Nach demselben Prinzip und mit Hilfe eines vordefinierten Flugplans kann nun eine autonome Navigation realisiert werden. Vielfach wird heute für die Bestimmung des Fluglagezustandes von Mikro- und Mini-UA-Plattformen auf die KF-basierte Datenfusion zurückgegriffen (vgl. Abschnitt 2.4). Dieses Verfahren verwendet hochfrequent zur Verfügung stehende Beobachtungen (hier aus einer inertialen Messeinheit) aus denen der gesuchte Systemzustand nach dem Prinzip der Koppelnavigation abgeleitet wird. Aus sich aufkumulierenden Messfehlern in den Beobachtungen entsteht eine zeitlich anwachsende Abweichung zum wahren Systemzustand. Die auftretende Abweichung ist unerwünscht und kann vermieden werden, indem zusätzliche tieffrequente unabhängige mit GNSS ermittelte Positionsupdates einbezogen werden, welche die Navigationslösung stützen. Zusätzlich werden vielfach Barometer und Magnetkompassensoren mitintegriert. Mit dem Barometer wird die mit GNSS schlecht bestimmte Höhenkomponente mit beobachtet und die Magnetkompassensoren übernehmen die Stützung der Navigationslösung während GNSS-Signalausfällen. Entsprechende Realisierungen von Fluglagekontrollern, speziell entwickelt für Mini- und Mikroplattformen, werden in Winkler (2007), Wendel et al. (2006) oder Eck (2001) ausführlich beschrieben und diskutiert.

2.5.3 INS/GNSS-basierte Navigationssysteme

Für die Bestimmung der Fluglage von Mikro- und Minidrohnen (vgl. Abschnitt 2.5.2) oder die direkte Messung der Bildorientierung, beispielsweise in der Luftbildphotogrammetrie (vgl. Abschnitt 2.3.2), werden INS/GNSS-basierte Navigationssysteme verwendet. Grundprinzip der integrierten Navigation ist die Kombination unterschiedlicher Sensoren, deren Messdaten sich gegenseitig kontrollieren oder ergänzen und damit die Navigationslösung stützen. Winkler (2007) zeigt die komplementären Eigenschaften der Inertialmesstechnik und der Satellitennavigation für die Schätzung der Trägerplattformposition und -ausrichtung.

Eigenschaft	INS	GNSS
Erfassung hochdynamischer Bewegungen	ja	eingeschränkt
Langzeitgenauigkeit	gering	hoch
Störbarkeit von aussen	nein	ja
Zuverlässigkeit	hoch	eingeschränkt

Tabelle 2-4: Komplementäre Eigenschaften INS vs. GNSS

Die Kombination von INS- und GNSS-Sensoren ist gerade dank der möglichen Miniaturisierung beider Sensoren speziell für den Einsatz auf Mikro- und

Miniplattformen geeignet. Nachfolgend wird kurz auf die beiden Sensortypen eingegangen.

2.5.3.1 Inertiale Navigationssysteme

Die inertielle Messtechnik oder Trägheitsnavigation macht sich die Massenträgheit von Körpern gegenüber linearen Beschleunigungen und Drehbeschleunigungen zu Nutze. Für die Bestimmung der Verschiebung sowie der Ausrichtung im Raum werden in einer inertialen Messeinheit (IMU - engl. *Inertial Measurement Unit*) drei Beschleunigungssensoren sowie drei Drehratensensoren (Gyros) verwendet, die je auf drei normal zueinander stehenden Achsen angebracht werden. Die Aufgabe der inertialen Messeinheit besteht darin, die linearen Beschleunigungen sowie die Drehraten der Gyros aufzuzeichnen. Das inertielle Navigationssystem bestimmt nun mit Hilfe der Navigationsrechnung (Jekeli, 2001) aus den aufgezeichneten Rohmesswerten der IMU die aktuelle Position und Ausrichtung im Raum. Inertiale Navigationssysteme werden unterteilt in Plattformsysteme (Kardanische Aufhängung) und trägerorientierte Systeme (engl. *Strap-down Systems*), die sich hinsichtlich der Anordnung der einzelnen Sensoren und ihrer mechanischen Realisierung unterscheiden. Inertiale Navigationssysteme werden heute in unterschiedlichen Güte- und Preisklassen produziert.

	Systemgenauigkeit			
	hoch <i>strategic grade</i>	mittel <i>navigation grade</i>	niedrig <i>tactical grade</i>	tief <i>MEMS-type</i>
Position [m]				
Nach 1h	300-500	$1 \cdot 3 \cdot 10^3$	$200-300 \cdot 10^3$	-
Nach 1sec	0.01-0.02	0.03-0.1	0.3-0.5	-
Ausrichtung [°]				
Nach 1h	$3-8 \cdot 10^{-3}$	0.01-0.05	1-3	>100
Nach 1sec	$< 0.3 \cdot 10^{-3}$	$3-5 \cdot 10^{-3}$	0.01-0.03	~1
Preis [€]	~1'000'000	~100'000	~10'000	<100

Tabelle 2-5: Klassen inertialer Navigationssysteme

Tabelle 2-5 zeigt eine Klassifizierung der Systeme nach Schwarz (1995) und Wägli (2009). Zu beachten gilt, dass auf Mini- und Mikroplattformen hauptsächlich MEMS- (*Micro-Electro-Mechanical Systems*) basierte INS (Titterton und Weston, 2004, Kap. 7) zum Einsatz kommen, da diese sehr klein, leicht und preisgünstig sind. Entsprechend ist die Güte dieser Systeme verglichen mit *navigation grade* und *tactical grade* INS, wie sie für Kartierungsanwendungen in der luftgestützten Photogrammetrie heute Verwendung finden, stark reduziert. Für weiterführende Grundlagen zur inertialen Messtechnik sei beispielsweise auf Jekeli (2001) verwiesen.

2.5.3.2 Globale Navigationssatellitensysteme (GNSS)

Navigationssatellitenempfänger werden heute für verschiedenste Anwendungen eingesetzt und unterscheiden sich stark hinsichtlich der implementierten Auswerteverfahren und der daraus resultierenden Güte. Nachfolgende Tabelle gibt einen Überblick über die am Markt erhältlichen GPS-Empfängertypen hinsichtlich der erreichbaren Positionierungsgenauigkeit, implementierten Auswerteverfahren und Preiskategorien.

	Navigationsempfänger			Geodätische Empfänger	
	1	1	1	2	2
Frequenzen	1	1	1	2	2
Beobachtungstypen	Code	Code	Code, Trägerphase	Code, Trägerphase	Code, Trägerphase
Auswerteverfahren	Pseudorange	differentiell (DGPS)	phasengeglättet	Precise Point Positioning	differentiell KAR,RTK
Genauigkeit					
Lage [m]	5-10	2-5	0.3-2	0.1	0.01-0.02
Höhe [m]	10-15	4-8	2	0.2	0.02-0.04
Preis [€]	1-100	1-100	1000	10k-20k	10k-20k

Tabelle 2-6: Klassen Satellitennavigationsempfänger für kinematische Anwendungen

Auf Mikro- und Miniplattformen kommen aus Gewichts- wie auch aus Kostengründen meist Navigationsempfänger zum Einsatz. Diese Empfänger-Kategorie wird dadurch charakterisiert, dass die Positionsschätzung aus sogenannten Code-Beobachtungen bzw. den gemessenen Pseudodistanzen (engl. *Pseudorange*) zwischen Empfänger und Satellit erfolgt. Viele Empfängertypen dieser Kategorie können zusätzlich im differentiellen Modus (Abk. DGPS) betrieben werden, indem zusätzliche Pseudorange Korrekturen einer Referenzstation mitverarbeitet werden. Zum Vergleich: für die direkte Georeferenzierung in der Luftbildphotogrammetrie werden geodätische Empfänger eingesetzt, mit denen um einen Faktor 10^2 höhere Genauigkeiten erreicht werden. Vertiefende Informationen über die Satellitennavigation allgemein und die Auswerteverfahren im Speziellen sind in der einschlägigen Fachliteratur beispielsweise in Leick (2004) zu finden.

2.5.4 Videotechnik und -übertragung

Ein Video besteht aus kontinuierlich aufgenommenen Einzelbildern, welche sich durch analoge oder digitale Technologie übertragen lassen. Für die analoge Fernsehsignalübertragung existieren unterschiedliche Standards. Grösstenteils wird in Westeuropa dafür PAL (engl. *Phase Alternating Line*) verwendet, welcher 25

Einzelbilder mit 625 Zeilen pro Sekunde übermittelt. Von den 625 übermittelten Zeilen sind 575 Zeilen sichtbar, was die Übermittlung zusätzlicher benutzerdefinierter Daten ermöglicht. Zeitstempel können analog nach den Verfahren VITC (engl. *Vertical Interval TimeCode*) und LTC (engl. *Longitudinal Timecode*) in das Videosignal integriert werden. Das Verfahren VITC führt die zusätzliche Zeitinformation zu jedem – Einzelbild – in einer nicht sichtbaren Zeile. Vielfach wird dafür die Zeile 21 verwendet. Das zweite Verfahren LTC übermittelt die Zeitinformation zu jedem Einzelbild wenn vorhanden in der zusätzlichen Tonspur. Im Gegensatz zur analogen Übertragung erfordert die digitale Verarbeitung eine vorgängige A/D-Wandlung des Videosignals. Dabei wird jedes Einzelframe in eine Matrixrepräsentation abgebildet. Damit die anfallende riesige Datenmenge übertragen und verarbeitet werden kann, werden effiziente Kompressionsverfahren benötigt. Unterschieden werden hier Ansätze, welche basierend auf bekannten Bildkompressionen auf Einzelbildbasis komprimieren sowie Verfahren, welche lediglich die Änderungen im Bild über die Videosequenz aufzeichnen, womit eine starke Reduktion der anfallenden Datenmenge erzielt wird. Aufgrund der involvierten Kompressions- und Dekompressionsalgorithmen muss bei der digitalen Videoübertragung eine Übertragungsverzögerung im Bereich mehrerer Frames in Kauf genommen werden. Für weiterführende Informationen zur analogen und digitalen Videotechnik ist auf das Grundlagewerk von (Schmidt, 2005) verwiesen.

3 Mobile Video- Georegistrierung

Der erste Abschnitt dieses Kapitels führt in das entworfene und implementierte Konzept zur Georegistrierung mobil erfasster Videosequenzen ein. Die beiden anschliessenden Abschnitte behandeln im Detail den hier vorgeschlagenen Video-Georegistrierungsansatz und die dafür benötigten Grundlagen. Für die weiterführenden Betrachtungen ist zwischen den Begriffen Georegistrierung und Georeferenzierung zu unterscheiden. Der Begriff Georegistrierung bezeichnet die optimale Aufnahmeausrichtung und -position der Videokamera bezüglich eines digital vorliegenden Geodatensatzes wie beispielsweise eines bestehenden 3D-Stadtmodells. Im Gegensatz dazu beschreibt der Begriff Georeferenzierung die Ausrichtungs- und Aufnahmetrajektorie der Videokamera bezüglich eines geodätisch definierten Bezugssystems. Die beiden Begriffe sind in diesem Falle äquivalent, wenn der digitale Geodatensatz fehlerfrei im selben geodätischen Bezugssystem vorliegt, und das Video-Sensormodell korrekt modelliert ist.

3.1 Architektur und Integrationskonzept

Das entworfene Video-Georegistrierungskonzept basiert auf den in Abschnitt 2.1.1 eingeführten photogrammetrischen Lösungsansätzen und ergänzt diese mit dem aus dem Umfeld der *Video Registration* vorgeschlagenen Konzept 3D-Modell-basierte Georegistrierung. Der hier vorgeschlagene Ansatz propagiert ein anderes Bild-zu-Modell-Zuordnungsverfahren als in Matai et al. (2003) vorgeschlagen und implementiert zusätzlich zur Stützung der Georegistrierungslösung Konzepte aus SLAM und SfM.

Einführend werden die für die Video-Georegistrierung vorausgesetzten Systemanforderungen besprochen. Nachfolgend wird die gesamte Echtzeit-Datenverarbeitungskette von der Mikro- und Minidrohnen gestützten Videoerfassung, über die Datenübertragung und Georegistrierung bis zur Integration und Nutzung der Daten in einem 3D-Geoinformationsdienst vorgestellt. Das vorgeschlagene Video-Georegistrierungskonzept wird im letzten Abschnitt dieses Kapitels eingeführt.

3.1.1 Ausgangslage und Voraussetzungen

Der anschliessend eingeführte Video-Georegistrierungsansatz setzt folgende Minimalanforderungen an die mobile Aufnahmeplattform und den Bildaufnahmesensor voraus. Die Aufnahmeplattform integriert eine Navigationseinheit zur Bestimmung des Fluglagezustandes, welcher die Plattformposition und -ausrichtung sowie die dazugehörige Zeitepoche enthält. Im Kontext mit unbemannten Flugplattformen wird hier vielfach von Fluglagedaten oder Telemetriedaten gesprochen.

$$\mathbf{X}_{WGS84} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{WGS84} = f(\varphi, \lambda, h_{ell})$$

Plattformposition im geozentrischen kartesischen Koordinatensystem (WGS 84). Die kartesische Koordinatenrepräsentation X, Y, Z kann mittels Transformation in die krummlinige Koordinatenrepräsentation geodätische Breite, Länge und ellipsoidische Höhe $\varphi, \lambda, h_{ell}$ und umgekehrt überführt werden (Torge, 2003).

$$\mathbf{R}_{nav}^{body} = (\omega, \theta, \psi)$$

Plattformausrichtung (Roll-, Nick- und Gierwinkel) bezüglich des Navigationskoordinatensystems (lokales Tangentialsystem). Details zur Parametrisierung der Rotationsmatrix sind in Anhang A zu finden.

$$t^{UTC}$$

Zeitepoche bezüglich UTC

Dabei stehen die Positions- und Ausrichtungsupdates mit einer Mindestfrequenz von 2 Hz zur Verfügung. Der Einfluss der verfügbaren Datenfrequenz wird später in dieser Arbeit im Detail diskutiert. Die Genauigkeit der Positions- und Ausrichtungsupdates muss in der Größenordnung liegen, die heute mit MEMS-INS/GNSS-Navigationssensoren typischerweise erreicht wird (vgl. Abschnitt 2.5.3 und Diskussion in Abschnitt 3.2.2.1, Position SEP 5-10m und Ausrichtung $<5^\circ$). Weiter wird angenommen, dass die Hebelarm- sowie die Fehlaustrichtungskomponenten des eingebauten Bildaufnahmesensors bezüglich des Plattformkörpersystems aus einer vorgängigen Kalibrierung bekannt sind (vgl. Abschnitt 2.3.2.1). Für die Bildaufnahme wird eine handelsübliche PAL-Videokamera vorausgesetzt, für welche die Parameter der inneren Orientierung x_0, y_0 und c ebenfalls aus einer vorgängigen Kalibrierung vorliegen. Im Weiteren ist vorausgesetzt, dass diese über die gesamte Videosequenz konstant bleiben. Für die Realisierung der integrierten Georegistrierung auf einen bestehenden Geodatenbestand ist zusätzlich eine Anwendung (z.B. virtueller Globus i3D, vgl. 2.5.1) erforderlich, die in Echtzeit, aufgrund eines definierten Sichtkegels die enthaltenen 3D-Objektgeometrien mit Semantik zurückliefert. Der für die Abfrage benötigte *View Frustum* ist durch die Projektionsmatrix festgelegt, welche sich aus der äusseren Orientierung (Aufnahmeposition und Ausrichtung bezüglich des übergeordneten Koordinatensystems) und der inneren Orientierung zusammensetzt.

3.1.2 Datenverarbeitungskette und Architektur

Die nachfolgende Abbildung 3-1 zeigt die Architektur der gesamten Verarbeitungskette von der Videodatenerfassung, über die Datenübertragung und die der Video-Georegistrierung bis hin zur Videointegration im virtuellen Globus i3D. Die untenstehenden Ausführungen zur Datenverarbeitung beziehen sich auf den für diese Arbeit angenommenen mini- und mikrodrohnenbasierten Anwendungsfall. Die integrierte Navigationskomponente auf der Aufnahmeplattform setzt keine spezielle Sensorkombination voraus und kann grundsätzlich mit unterschiedlichen Sensoren realisiert werden. Die integrierte Datenübertragungskomponente kann aus der

Architektur entfernt werden, wenn die Video-Georegistrierung und Datenverarbeitung direkt auf der Aufnahmeplattform erfolgen soll.

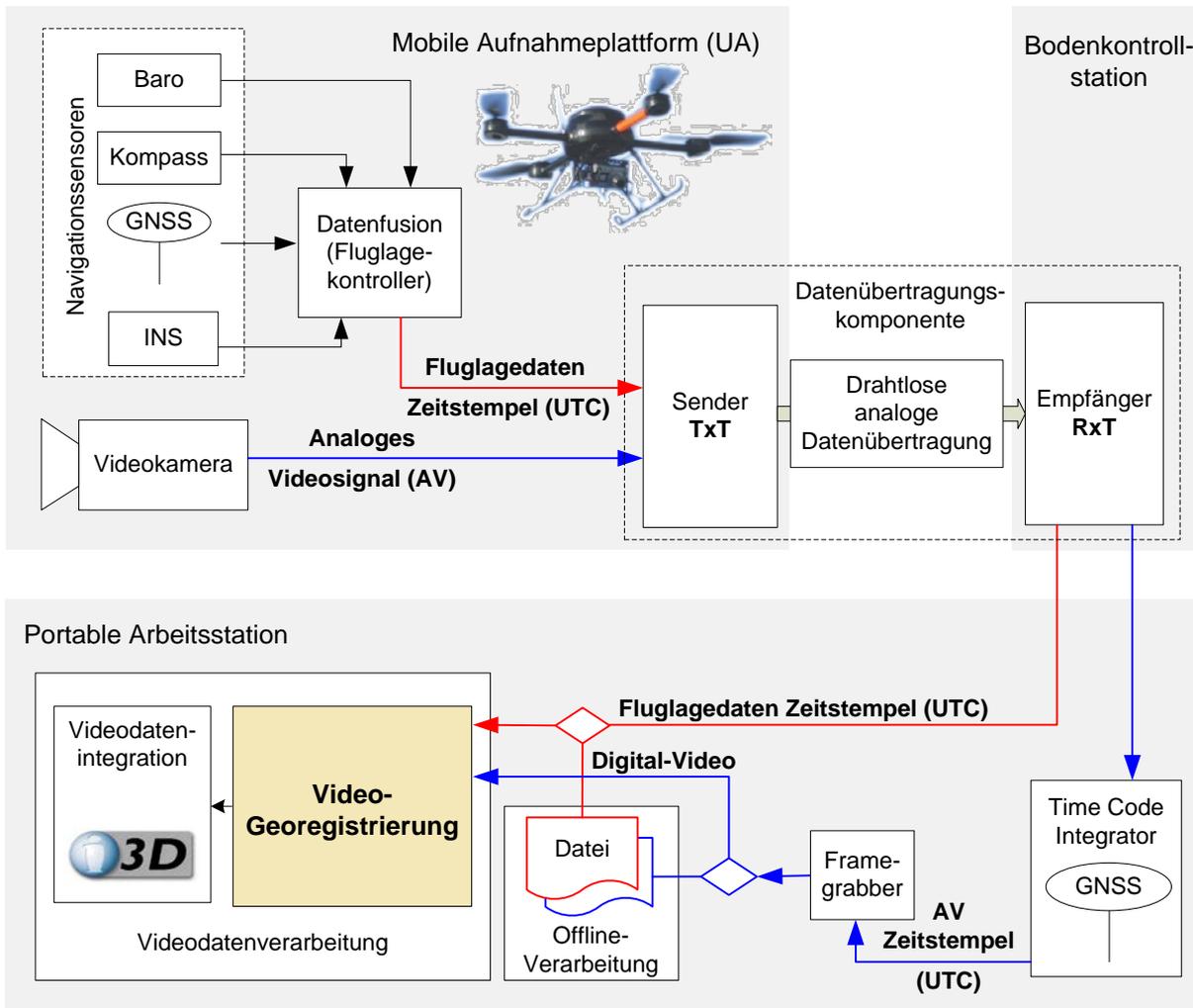


Abbildung 3-1: Videoverarbeitungskette: Video- und Fluglagedatenerfassung, Datenübertragung, Georegistrierung mit Videodatenintegration

Am Beginn der Videoverarbeitungskette steht die unbemannte Aufnahmeplattform (Mini- oder Mikrodrohne), die mit Fluglagekontroller, Videokamera und analogem Datenlinksender ausgerüstet ist. Der Fluglagekontroller liefert Fluglagenzustände bestehend aus Zeitstempel, Position, Ausrichtung und eventuell Plattformgeschwindigkeit mit einer Update-Frequenz von mindestens 2 Hz. Das von der Videokamera erfasste analoge Videosignal wird anschließend zusammen mit den Fluglagenzuständen über einen analogen Datenlink zur Bodenkontrollstation übertragen. Die an der Bodenkontrollstation empfangenen Fluglagenzustände und das Videosignal werden über getrennte Schnittstellen für die anschließende Verarbeitung zur Verfügung gestellt. Damit die erfassten Videodaten später mit Hilfe der aufgezeichneten Fluglagedaten georegistriert werden können, wird dem Videosignal mit Hilfe der Time-Code-Integrator-Komponente eine Zeitstempelinformation hinzugefügt. Diese Time-Code-Komponente wird mit dem Zeitsignal, generiert durch einen GNSS-Empfänger,

synchronisiert. Damit ist gewährleistet, dass die Zeitstempel der Fluglagezustände sowie die integrierten Zeitstempel der Videodaten in der gleichen Zeitreferenz vorliegen (beide UTC). Im Anschluss wird das Videosignal analog/digital gewandelt. Bei der Echtzeit-Videodatenintegration werden die beiden verfügbaren Datenströme direkt von einer portablen Arbeitsstation empfangen und mit den entsprechenden Video-Softwarekomponenten verarbeitet. Bei einer Offline-Anwendung können die beiden Datenströme auf einen sekundären Datenspeicher aufgezeichnet und zu einem späteren Zeitpunkt mit den Video-Softwarekomponenten verarbeitet werden. Damit die Zeitsynchronisation zwischen den Fluglagezuständen und dem analog/digital gewandeltem Video anschliessend bei der Video-Georegistrierung korrekt durchgeführt werden kann, muss in der aufgezeigten Architektur für die Datenübertragung ein analoger Datenlink eingesetzt werden. Bei dieser Übertragungstechnik resultiert im Gegensatz zur digitalen Übertragungstechnik eine vernachlässigbare kleine Laufzeitverzögerung

$$\Delta t = \frac{r}{c} \quad (10)$$

bestimmt aus der Datenübertragungreichweite r und der Lichtgeschwindigkeit c (vgl. Abschnitt 2.5.4). Wird eine digitale Datenübertragung verwendet, muss die Zeitstempelintegration vor der Datenkompression für die Datenübertragung auf der Plattform stattfinden oder die entstehende Zeitverzögerung Δt durch die Datenkomprimierung und -dekomprimierung muss konstant und mit ausreichender Genauigkeit vorliegen.

Mit Hilfe der nachfolgenden Videoverarbeitungs-komponenten kann die Videosequenz in einem ersten Schritt georegistriert und anschliessend in den virtuellen Globus i3D der Fachhochschule Nordwestschweiz integriert werden. Die Architektur der Video-Georegistrierung ist im nachfolgenden Abschnitt im Detail behandelt. Weitere Informationen zur Systemarchitektur sowie zu den einzelnen im virtuellen Globus i3D realisierten Video-Integrationsvarianten sind in Eugster und Nebiker (2008) zu finden.

3.1.3 Zweistufige Video-Georegistrierung

Für die eigentliche Video-Georegistrierung wird ein Konzept propagiert, das aus zwei nacheinander ausführbaren Registrierungsschritten besteht. Die erste Registrierungsstufe implementiert das aus der Photogrammetrie bekannte und etablierte auf Videosequenzen adaptierte Verfahren der direkten Georeferenzierung. Soll die Georegistrierungsgenauigkeit der direkt gemessenen äusseren Orientierung des Videos zusätzlich gesteigert werden, kann die zweite Registrierungsstufe in die Datenverarbeitung integriert werden. Die zweite Registrierungsstufe basiert auf der Idee der ebenfalls aus der Photogrammetrie bekannten integrierten oder kombinierten Georeferenzierung und nutzt zur Steigerung der Genauigkeit die verfügbaren 3D-Geometrielemente als Referenzgeometrie.

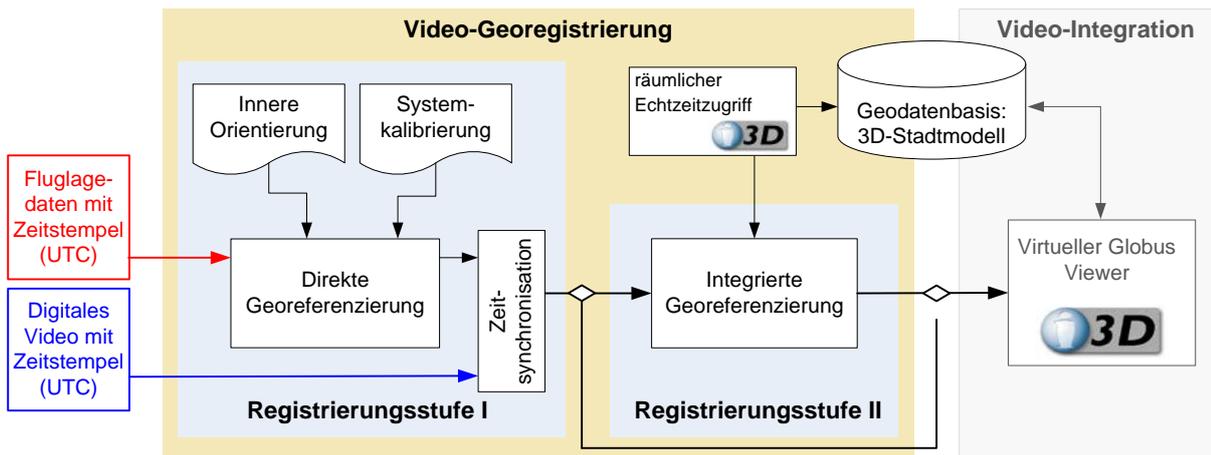


Abbildung 3-2: Architektur Video-Georegistrierung

Die Architektur des vorgeschlagenen zweistufigen Video-Georegistrierungsansatz ist in Abbildung 3-2 dargestellt. Werden auf der Aufnahmeplattform für die direkte Messung der Plattformposition und -ausrichtung Sensoren hoher Qualität verwendet oder genügt die erreichbare Genauigkeit der Sensoren geringer Qualität den Anforderungen eines bestimmten Anwendungsfalls, erübrigt sich die Verwendung der zweiten Registrierungsstufe in der Datenverarbeitung. Die zweistufige Georegistrierung bietet ein sehr flexibel anwendbares, mit unterschiedlichen Aufnahmeplattformen und Sensoren kombinierbares Konzept.

Die Verwendung der direkten Georeferenzierung als Grundlage für die Video-Georegistrierung kann in mehrfacher Hinsicht begründet werden. Videodaten weisen einen visuell kontinuierlichen Charakter auf, was für die Georegistrierung implizit eine kontinuierlich verfügbare äussere Orientierung verlangt. Inertiale Messeinheiten zeichnen sich speziell durch die Eigenschaft aus, dass die dafür verwendeten Beschleunigungs- und Kreiselsensoren sehr hohe Beobachtungsraten erreichen und damit eine Koppelnavigationslösung von typischerweise 100-200 Hz realisiert werden kann. Grundsätzlich können damit die mit einer viel geringeren Frequenz vorliegenden Einzelbilder (25 - 30 Hz) einer Videosequenz optimal durch Interpolation georeferenziert werden. Zudem kann die Lösung der mit Hilfe der Inertialsensoren realisierten Koppelnavigation durch tiefer-frequente Updates (z.B. 1 Hz) beispielsweise durch GNSS oder ein visionbasiertes System gestützt werden. Im Gegensatz zu den Ansätzen, welche die Beziehung zwischen erfasster Umgebung und Aufnahmesystem nur mit Hilfe von Bildaufnahmen lösen (vgl. Tabelle 2-1), muss bei der Nutzung der direkten Georeferenzierung nicht jedes Bild indirekt orientiert werden, um für jedes Bild die gesuchte Aufnahmeposition und -ausrichtung abzuleiten. Damit lässt sich der Berechnungsaufwand entscheidend reduzieren und in nahezu statischen Aufnahme-konfigurationen entstehen keine Singularitäts- oder Initialisierungsprobleme. Ein weiterer Vorteil der Nutzung der Navigationsdaten aus INS/GNSS ist der direkte Bezug zu einem übergeordneten globalen geodätischen Referenzsystem. Ein weiterer entscheidender Faktor im dargestellten Konzept ist die Nutzung einer *Terrain Engine* für die Abfrage von 3D-Stadtmodellldaten für die integrierte Georeferenzierung. Diese

Technologie ermöglicht im Gegensatz zu Geografischen Informationssystemen (GIS) aus folgenden Gründen einen Echt- oder Nahechtzeitzugriff auf einen Geodatenbestand. Virtuelle Globen implementieren hoch effiziente räumliche Datenstrukturen zur Visualisierung der 3D-Geodatenbasis. Spezialisierte *Streaming*-Algorithmen (fortlaufendes Nachladen der Geodaten) basierend auf räumlichen Datenstrukturen ermöglichen eine kontinuierliche Aktualisierung der Geodatenbasis im Primärspeicher. Wird nun die aktuelle Plattformposition für die Aktualisierung bzw. für das *Streaming* verwendet, lässt sich mit Hilfe dieser Technologie ein effizienter Datenzugriff realisieren (vgl. Abschnitt 2.5.1).

3.2 Direkte Video-Georeferenzierung - Registrierungsstufe I

Die erste Registrierungsstufe verwendet den in Abschnitt 3.1.3 bereits eingeführten Ansatz der direkten Georeferenzierung. Dieses Kapitel behandelt auf der einen Seite wie der etablierte und bekannte Ansatz für die Georeferenzierung mobil erfasster Videosequenzen genutzt werden kann. Auf der anderen Seite werden die Konsequenzen diskutiert, welche der Einsatz von Navigationssensoren geringer Qualität auf die resultierende direkte Georeferenzierungslösung hat. Entsprechende Schlussfolgerungen aus diesen Untersuchungen werden im Konzept der zweiten Registrierungsstufe berücksichtigt und begründen die vorgeschlagene integrierte Lösung.

3.2.1 Umsetzung direkte Video-Georeferenzierung

Im Kontext der Video-Georegistrierung formuliert, wird diejenige Videoaufnahme-position- und Orientierung gesucht, an der eine aus der Videosequenz erfasste 3D-Geometrie mit einer bekannten 3D-Geometrie aus dem übergeordneten Referenzrahmen zusammenfällt. Im Zusammenhang mit der Registrierung der erfassten Videodaten mit 3D-Geometriedaten aus dem virtuellen Globus muss die gesuchte Positions- und Orientierungsinformation der Aufnahmetrajektorie bezüglich eines global gelagerten geodätischen Referenzsystems (hier WGS 84) georeferenziert werden. Damit ist gewährleistet, dass die georeferenzierten Videodaten direkt in den virtuellen Globus integriert und die erreichte Georegistrierungsgenauigkeit überprüft werden kann. Nachfolgend wird die eingeführte direkte Georeferenzierung (vgl. Abschnitt 2.3.2) auf mittels Mini- und Mikrodrohnen erfasste Videosequenzen adaptiert, bevor die Video- und Fluglagedatenverarbeitung der ersten Registrierungsstufe im Detail vorgestellt wird.

3.2.1.1 Direkte Georeferenzierung mit Mini- und Mikrodrohnen

Im Vergleich zu bemannten flugzeuggestützten photogrammetrischen Bfliegungen unterscheidet sich die Ausgangslage für mit Mini- und Mikrodrohnen durchgeführte Kartierungsmissionen wesentlich. Aus Gewichts- wie auch aus Kostengründen kann für die Georeferenzierung nicht auf ein INS/GNSS-Navigationssystem hoher Qualität zurückgegriffen werden. Für die direkte Georeferenzierung können damit nur die bereits für die autonome Navigation und Stabilisierung integrierten Navigationssensoren geringer Qualität verwendet werden. Vielfach werden die entsprechenden Video- oder Bildkameras an einer sogenannten kardanischen Aufhängung (engl. *Gimbal*) an die Trägerplattform montiert. Diese Aufhängungen erlauben die unabhängige Ausrichtung der Kamera bezüglich des Plattformkoordinatensystems und integrieren vielfach eine automatische Bildstabilisation.

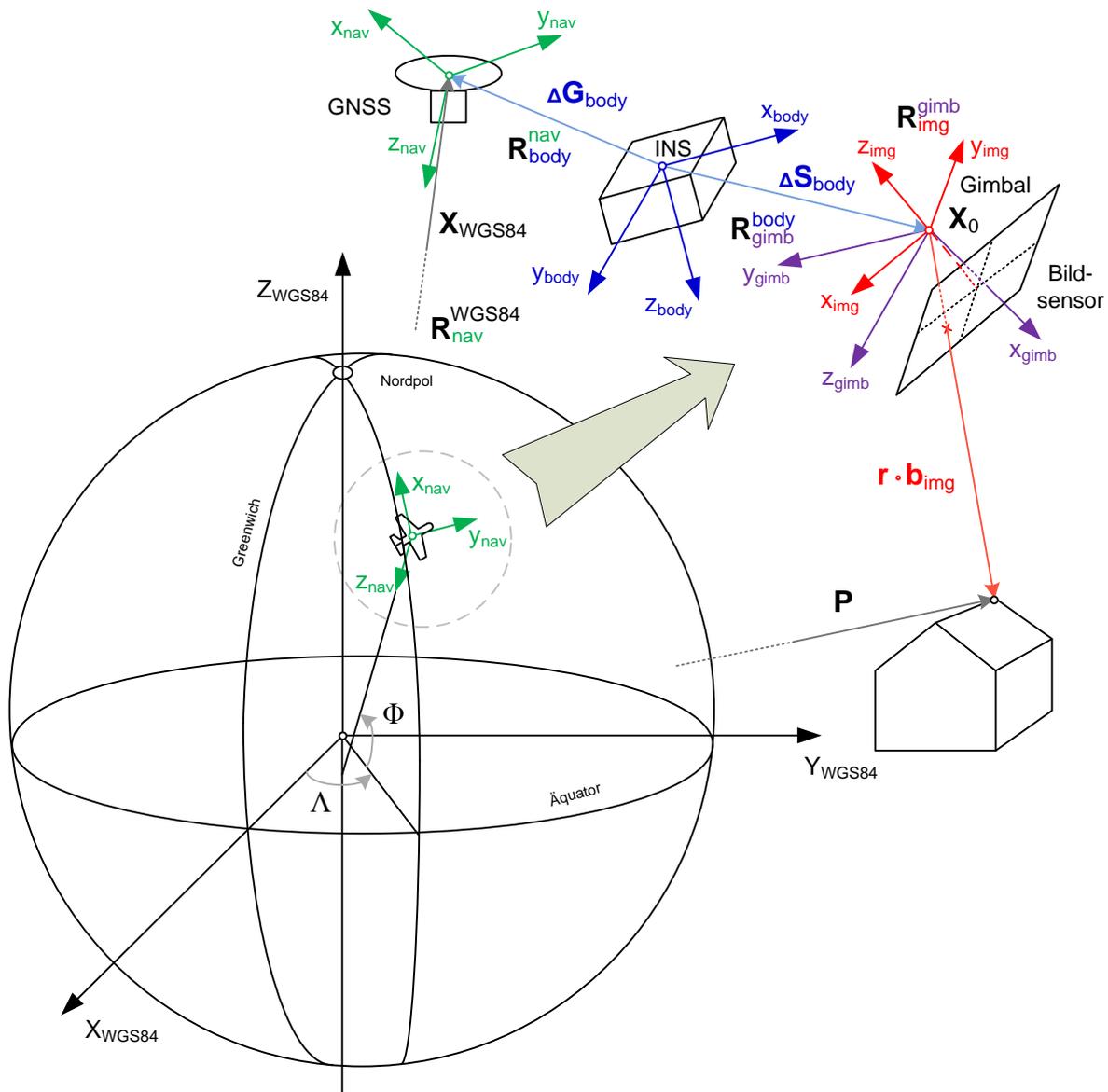


Abbildung 3-3: Aufnahmegeometrie der mini- und mikrodrohnenbasierten direkten Georeferenzierung

Bedingt durch den Einsatz steuerbarer Kameraaufhängungen kann die benötigte rigide Beziehung zwischen INS und Kamerakoordinatensystem nicht direkt erhalten werden. Aufgrund der kleinen Trägerplattformen entstehen sehr kleine Hebelarmkomponenten (<20cm), die meist aufgrund der erreichbaren Navigationslösungsgenauigkeit vernachlässigt werden können. Abbildung 3-3 verdeutlicht die angepasste Ausgangslage. Wird diese entsprechend berücksichtigt, kann die direkte Georeferenzierung (vgl. Gl. (C-5)) wie folgt ausformuliert werden. Die allgemein formulierte Drehung vom Plattformsystem ins übergeordnete Koordinatensystem kann durch die beiden Teilrotationen

$$\mathbf{R}_{body}^{WGS84} = \mathbf{R}_{nav}^{WGS84} \cdot \mathbf{R}_{body}^{nav} \quad (11)$$

mit

$\mathbf{R}_{nav}^{WGS84} = R(\varphi, \lambda)$ Rotationsmatrix welche die Drehung vom Navigationskoordinatensystem ins geodätische Referenzsystem (WGS 84) mittels geodätischer Breite und Länge der aktuellen Plattformposition ausdrückt. (Details zur Parametrisierung der Rotationsmatrix sind in Anhang A zu finden)

$\mathbf{R}_{body}^{nav} = R(\omega, \theta, \psi)$ Rotationsmatrix welche die Drehung vom INS-Körpersystem ins Navigationssystem beschreibt. (vgl. Kap. 3.1.1)

ausgedrückt werden und beschreibt die Verdrehung des INS-Körpersystems (Plattformkoordinatensystem) zum global gelagerten geodätischen Referenzsystem WGS 84. Die bis anhin als FehlAusrichtung bezeichnete Drehung vom Bildkoordinatensystem ins INS-Körperkoordinatensystem kann aufgrund der erwähnten beweglichen Kameraaufhängungen wie folgt

$$\mathbf{R}_{sen}^{body} = \mathbf{R}_{gimb}^{body} \cdot \mathbf{R}_{img}^{gimb} \quad (12)$$

durch die beiden Rotationen

$\mathbf{R}_{gimb}^{body} = R(\alpha, \nu)$ Beschreibung der aktuellen Ausrichtung der Kameraaufhängung bezüglich INS-Körpersystem. (Details zur Parametrisierung der Rotationsmatrix sind in Anhang A zu finden)

$\mathbf{R}_{img}^{gimb} = R(\delta\omega, \delta\theta, \delta\psi)$ Rotationsmatrix welche die verbleibende FehlAusrichtung beschreibt. (Details zur Parametrisierung der Rotationsmatrix sind in Anhang A zu finden)

ersetzt werden. Was auf folgende adaptierte Gleichung für die direkte Georeferenzierung führt:

$$\mathbf{P} = \mathbf{X}_{WGS84} + \mathbf{R}_{nav}^{WGS84} \cdot \mathbf{R}_{body}^{nav} \left(r \cdot \mathbf{R}_{gimb}^{body} \cdot \mathbf{R}_{img}^{gimb} \cdot \mathbf{b}_{img} + \Delta \mathbf{S}_{body} - \Delta \mathbf{G}_{body} \right) \quad (13)$$

Folgende Substitutionen

$$\mathbf{R}_{img}^{world} = \mathbf{R}_{nav}^{WGS84} \cdot \mathbf{R}_{body}^{nav} \cdot \mathbf{R}_{gimb}^{body} \cdot \mathbf{R}_{img}^{gimb} \quad (14)$$

$$\mathbf{R}_{body}^{WGS84} = \mathbf{R}_{nav}^{WGS84} \cdot \mathbf{R}_{body}^{nav} \quad (15)$$

$$\Delta \mathbf{H}_{body} = \Delta \mathbf{S}_{body} - \Delta \mathbf{G}_{body} \quad (16)$$

führen auf

$$\mathbf{P} = \mathbf{X}_{WGS84} + \mathbf{R}_{body}^{WGS84} \cdot \Delta \mathbf{H}_{body} + \lambda \cdot \mathbf{R}_{img}^{world} \cdot \mathbf{b}_{img} \quad (17)$$

mit welcher aus mindestens zwei Bildern unbekannte Objektpunkte mittels Vorwärtsschnitt im globalen geodätischen Referenzsystem berechnet werden können. Wird die erhaltene Gl. (17) wie folgt nach

$$\mathbf{b}_{img} = \frac{1}{\lambda} \cdot \mathbf{R}_{world}^{img} (\mathbf{P} - \mathbf{X}_{WGS84} - \mathbf{R}_{body}^{WGS84} \cdot \Delta \mathbf{H}_{body}) \quad (18)$$

mit

$$\mathbf{R}_{world}^{img} = (\mathbf{R}_{img}^{world})^{-1} = (\mathbf{R}_{img}^{world})^T \quad (19)$$

aufgelöst, resultiert die bekannte Beziehung der perspektivischen Abbildung. Damit lassen sich beliebige bekannte Objektpunkte (z.B. aus bestehenden 3D-Stadtmodellen) im globalen geodätischen Referenzsystem in die Bildebene abbilden. Werden zusätzlich die beiden Hebelarm-Translationen berücksichtigt

$$\mathbf{X}_0 = \mathbf{X}_{WGS84} + \mathbf{R}_{body}^{WGS84} \cdot \Delta \mathbf{H}_{body} \quad (20)$$

resultiert direkt das Projektionszentrum \mathbf{X}_0 des Aufnahmesensors. Damit vereinfachen sich die beiden Beziehungen (17) und (18) zu

$$\mathbf{P} = \mathbf{X}_0 + \lambda \cdot \mathbf{R}_{img}^{world} \cdot \mathbf{b}_{img} \quad (21)$$

$$\mathbf{b}_{img} = \frac{1}{\lambda} \cdot \mathbf{R}_{world}^{img} (\mathbf{P} - \mathbf{X}_0) \quad (22)$$

aus welchen die Kollinearitätsgleichungen (C-2) resultieren. Damit ist die direkte Georegistrierung (oder AR-Lösung) beliebiger 3D-Stadtmodellldaten mittels von Mini- oder Mikrodrohnen erfasster Bild- und Navigationsdaten aufgezeigt.

3.2.1.2 Navigationsdaten- und Videoverarbeitung und Synchronisation

Ziel der ersten Registrierungsstufe ist die Berechnung des korrekten Sensormodells für jedes Einzelbild der Videosequenz. Das Sensormodell besteht aus den Parametern der äusseren und inneren Orientierung der Videokamera. Mit der vorliegenden Kenntnis des Sensormodells ist die Videosequenz bezüglich des übergeordneten Koordinaten-

systems georeferenziert. Die Architektur der dafür benötigten Video- und Navigationsdatenverarbeitung ist in Abbildung 3-4 dargestellt.

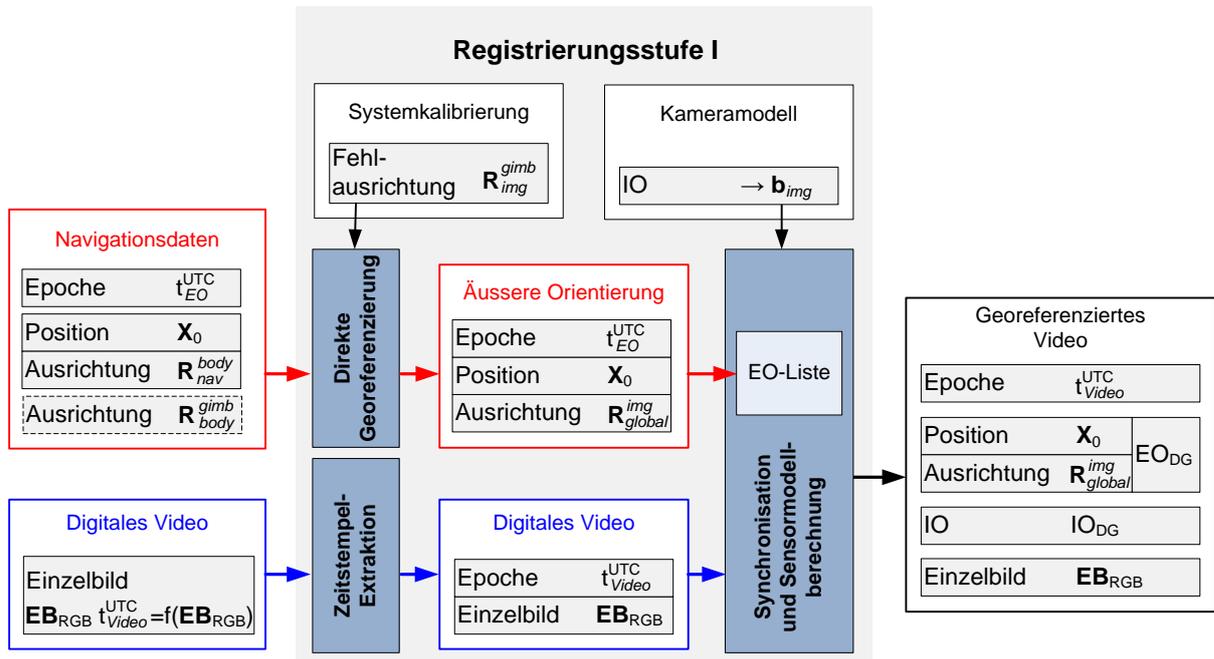


Abbildung 3-4: Architektur Registrierungsstufe I - Direkte Georeferenzierung

Ausgangslage für die Verarbeitung bilden die A/D gewandelte Videosequenz sowie die parallel zur Verfügung stehenden Navigationsdaten der Aufnahmeplattform. Zu beachten ist, dass Navigationsdaten Zustände meist in einer viel tieferen Frequenz übertragen werden. Im Vergleich zu den Videodaten (25-30 Hz) stehen diese Daten für Echtzeitanwendungen typischerweise lediglich mit $\sim 2-5$ Hz zur Verfügung. Diese ist abhängig vom eingesetzten Fluglagekontroller und der eingesetzten Datenlink-Komponente zwischen Aufnahmeplattform und Bodenk Kontrollstation. Wird auf der Aufnahmeplattform zusätzlich eine bewegliche Kameraaufhängung eingesetzt, muss die aktuelle Ausrichtung der Aufhängung bezüglich des Körperkoordinatensystems abgegriffen und mit den Navigationsdaten übermittelt werden. Vielfach werden die Plattformposition \mathbf{X}_{WGS84} , die Plattformausrichtung ω, θ, ψ und Kameraausrichtung α, ν in einer unterschiedlichen Frequenz abgetastet und übermittelt. Für nachfolgende Betrachtungen wird davon ausgegangen, dass diese Daten mit der gleichen Frequenz bezüglich UTC vorliegen. Ist dies nicht der Fall, können die Daten durch vorgängig implementierte Interpolationsverfahren auf die Ausgangslage wie in Abbildung 3-4 dargestellt gebracht werden. Zusätzlich zu den beiden Datenströmen werden die für die direkte Georeferenzierung benötigten kalibrierten Fehl-ausrichtungswinkel $\delta\omega, \delta\theta, \delta\psi$ sowie die Parameter der inneren Orientierung \mathbf{IO} der verwendeten Videokamera als bekannt vorausgesetzt. In der Komponente 'Direkte Georeferenzierung' werden nun mit den zur Verfügung stehenden Navigationsdaten und der bekannten Fehl-ausrichtung für jede Epoche t_{EO}^{UTC} nach Gl. (14) und (20) die Komponenten der äusseren Orientierung (\mathbf{X}_0 und \mathbf{R}_{img}^{world}) der Videokamera berechnet. Die resultierenden äusseren

Orientierungszustände werden anschliessend in der Komponente 'Synchronisation und Sensormodellberechnung' in einer Liste für die weitere Verarbeitung zwischengespeichert. Parallel dazu, wird in der Komponente 'Zeitstempel-Extraktion' der in jedem Einzelbild der Videosequenz integrierte Zeitstempel extrahiert. Die Epoche t_{Video}^{UTC} liegt nach der Zeitstempel-Integration (vgl. Abbildung 3-1) und A/D-Wandlung als Binär-Code in einer Bildzeile jedes Einzelbildes \mathbf{EB}_{RGB} vor. Mittels Kreuzkorrelation kann die implizit enthaltene Zeitstempelinformation t_{Video}^{UTC} aus jedem Einzelbild \mathbf{EB}_{RGB} extrahiert werden. Damit stehen sämtliche Informationen zur Verfügung, die in der Komponente 'Synchronisation und Sensormodellberechnung' benötigt werden, um für jedes Einzelbild die gesuchte direkt gemessene Sensororientierung zu berechnen. Für jede Epoche t_{Video}^{UTC} kann die Kameraposition und Ausrichtung wie folgt

$$\begin{pmatrix} \mathbf{X}_0 \\ \Theta \end{pmatrix}_{DG}^{t_{Video}^{UTC}} = \begin{pmatrix} \mathbf{X}_0 \\ \Theta \end{pmatrix}_i + \frac{\Delta t}{\Delta T_{EO}} \left(\begin{pmatrix} \mathbf{X}_0 \\ \Theta \end{pmatrix}_i - \begin{pmatrix} \mathbf{X}_0 \\ \Theta \end{pmatrix}_{i-1} \right) \quad (23)$$

aus den zwischengespeicherten äusseren Orientierungen bestimmt werden. Dabei bezeichnen:

$$\Delta t = t_{Video}^{UTC} - t_{EO\ i}^{UTC}$$

Zeitdifferenz zwischen Epoche des Einzelbildes aus der Videosequenz und der aktuellsten zwischengespeicherten äusseren Orientierung.

$$\Delta T_{EO} = t_{EO\ i}^{UTC} - t_{EO\ i-1}^{UTC}$$

Zeitdifferenz zwischen der aktuellsten i und der zweit-aktuellsten $i-1$ zwischengespeicherten äusseren Orientierung.

$$\Theta = \begin{pmatrix} \omega \\ \varphi \\ \kappa \end{pmatrix} = f(\mathbf{R}_{world}^{img})$$

Eulerwinkel-Repräsentation, welche die Rotation zwischen dem globalen geodätischen Referenzsystem und dem Bildkoordinatensystem beschreibt. Für Details zur Konversion zwischen beiden Repräsentationen ist auf Craig (1986) verwiesen.

Es können folgende zwei Fälle unterschieden werden:

- $t_{Video}^{UTC} \leq t_{EO\ i}^{UTC}$:

Die Gleichung (23) beschreibt eine lineare Interpolation. Die Interpolation kommt ausschliesslich bei der nachträglichen Datenverarbeitung (*offline Processing*) zum Einsatz. Für diesen Fall kann gewährleistet werden, dass zu jedem Zeitpunkt eine aktuellere zwischengespeicherte äussere Orientierung verfügbar ist.

- $t_{Video}^{UTC} > t_{EO\ i}^{UTC}$:

Dieser Fall entspricht einer linearen Extrapolation aus den letzten beiden bekannten zur Verfügung stehenden äusseren Orientierungen. Im Allgemeinen wird dieser Fall bei der Echtzeit-Verarbeitung (*online Processing*) benötigt, da meist die Navigationsdaten durch diverse Verzögerungen später beim Synchronisationsmodul eintreffen.

Die zwischengespeicherten äusseren Orientierungen in der Komponente 'Synchronisation und Sensormodellberechnung' werden aus der Liste entfernt, sobald diese für die Inter- oder Extrapolation nicht mehr benötigt werden. Zu beachten ist in Gl. (23) die verwendete Eulerwinkel-Repräsentation für die Inter- und Extrapolation der Kameraausrichtung. Diese Repräsentation wird hier verwendet, da die Rotationsmatrix-Repräsentation nicht direkt elementweise inter- oder extrapoliert werden kann, ohne dass die Orthogonalität verloren geht. Nachteil bei der Verwendung der Eulerwinkel-Repräsentation ist, dass bei der Konvertierung $\mathbf{R} \rightarrow \mathbf{\Theta}$ in speziellen Konstellationen eine Singularität entstehen kann. Eine mögliche Lösung bietet die Beschreibung der Kameraausrichtung mit Hilfe der Quaternion-Repräsentation. Quaternionen können im Gegensatz zur Rotationsmatrix-Repräsentation mit dem Algorithmus von Shoemake (1985) interpoliert oder extrapoliert werden, womit eine Konvertierung in die Eulerwinkel-Repräsentation entfallen würde. Bei Verwendung von Quaternionen müsste konsequenterweise die in Abschnitt 3.3 beschriebene photogrammetrische Bildorientierung vollständig in dieser Repräsentation ausformuliert werden (Bayoud, 2006). Das gesuchte Sensormodell kann nun für die Epoche mit der bekannten inneren Orientierung komplementiert werden. Damit ist jedes Einzelbild \mathbf{EB}_{RGB} aus der Videosequenz georeferenziert und verfügt zusätzlich zur Epoche $t_{\text{Video}}^{\text{UTC}}$ über die inter- oder extrapolierte äussere Orientierung \mathbf{EO}_{DG} bestehend aus $\mathbf{EO}_{\text{DG}} = (\mathbf{X}_0, \mathbf{\Theta})^T$ sowie die dazugehörige innere Orientierung \mathbf{IO}_{DG} .

Abschliessend seien noch einige Bemerkungen zur verfügbaren Navigationszustand-Frequenz erwähnt. Damit die ausgeführte Plattformbewegung korrekt erfasst und bei der Georeferenzierung berücksichtigt werden kann, muss das Theorem von Nyquist-Shannon entsprechend eingehalten werden:

$$f_{\text{abtast}} = 2 \cdot f_{\text{maximal}} \quad (24)$$

mit

f_{abtast}	Minimale benötigte Abtastfrequenz
f_{maximal}	Maximale Frequenz der ausgeführten Plattformbewegung

Im Speziellen von der Fluglageregelung nicht kompensierte Plattformbewegungen, verursacht meist durch Wind, können sehr hochfrequent auftreten. Gerade Mikrodrohnen sind diesbezüglich sehr anfällig und erreichen nicht die benötigte Stabilität, um mit vorliegenden Navigationsdaten von 4-5 Hz die Georeferenzierung korrekt ableiten zu können. Gerade bei starker Windbeeinflussung ist das verletzte Abtasttheorem durch überlagerte hochfrequente Bewegungen sicht- und nachweisbar. Weiterführende Betrachtungen und Untersuchungen in diesem Zusammenhang sind im Abschnitt 5.2.3.2 zu finden.

3.2.2 Navigationsdaten geringer Qualität

Die Nutzung von Navigationssensoren geringer Qualität hat verschiedene Einflüsse und Konsequenzen auf die eingeführte direkte Georeferenzierung. Auf der einen Seite ist die erreichbare direkt gemessene Genauigkeit der äusseren Orientierung um einige Zehnerpotenzen geringer, als dies bei traditionell in der Luftbildphotogrammetrie eingesetztem Instrumentarium zu erwarten ist. Andererseits müssen insbesondere die in Abschnitt 2.3.2.1 zusammengefassten etablierten Ansätze zur Systemkalibrierung unter der angepassten Ausgangslage untersucht und beurteilt werden. Folgender Abschnitt gibt einen Überblick über das typische Fehlerverhalten von mittels MEMS-INS und GNSS-Navigationsempfängern bestimmten Fluglagedaten. Im Anschluss werden die Erkenntnisse im Zusammenhang mit der Systemkalibrierung diskutiert und ein möglicher Lösungsansatz für die Verbesserung der Georegistrierungsgenauigkeit vorgestellt.

3.2.2.1 Genauigkeiten und Fehlerhaushalt

Konkrete Genauigkeitsuntersuchungen von MEMS-INS/GNSS-basierten Navigationslösungen für die direkte Georeferenzierung sind rar. Die meisten diesbezüglichen Untersuchungen wurden im Zusammenhang mit implementierten Fluglageregelungssystemen in der Regelungstechnik durchgeführt. In Vörsmann et al. (2007), Winkler (2007) und Wendel et al. (2006) wird die Genauigkeitsfrage lediglich im Zusammenhang mit der automatischen Stabilisierung und autonomen Navigation diskutiert. Erste Untersuchungen und Entwicklungen zu dieser Fragestellung in der Geomatik verfolgen das Ziel, dass für mobile Kartierungssysteme versucht wird, die dafür benötigten sehr teuren *navigation grade* INS mit sehr viel kostengünstigeren MEMS-INS zu ersetzen (El-Sheimy, 2009). Für die Analyse der fusionierten Navigationslösung wurden die Trajektorien unterschiedlicher MEMS-INS/GNSS-basierter Autopiloten mit einer Referenztrajektorie verglichen, die mit einem geodätischem INS/GNSS-System (Applanix POSLV 210) bestimmt worden war. Die Navigationsdaten der untersuchten Autopiloten erreichen folgende Genauigkeiten:

	microdrones ¹ md4-200	UAVNavigation ² AP04
Position		
Lage (1σ)	2.7 m	4.7 m
Höhe (1σ)	3.6 m	5.0 m
Ausrichtung		
Roll und Nick (1σ)	1.3 °	1.5 °
Gier (1σ)	4.1 °	6.3 °

Tabelle 3-1: Genauigkeiten für fusionierte Navigationslösungen aus MEMS-INS/GNSS-Sensoren basierten Autopiloten

¹ microdrones GmbH: <http://www.microdrones.com/startseite.php>

² UAVNavigation: <http://www.uavnavigation.com>

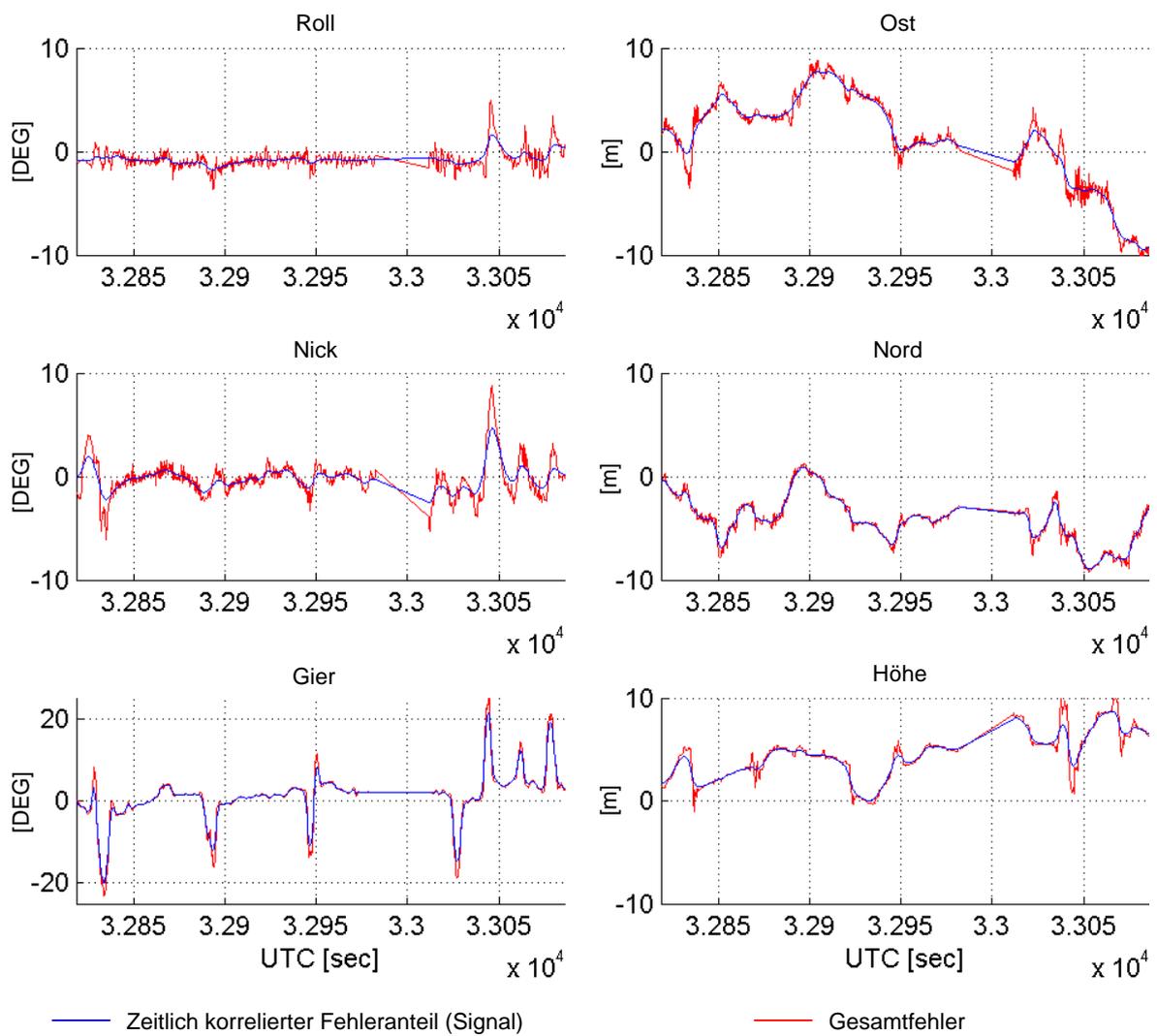
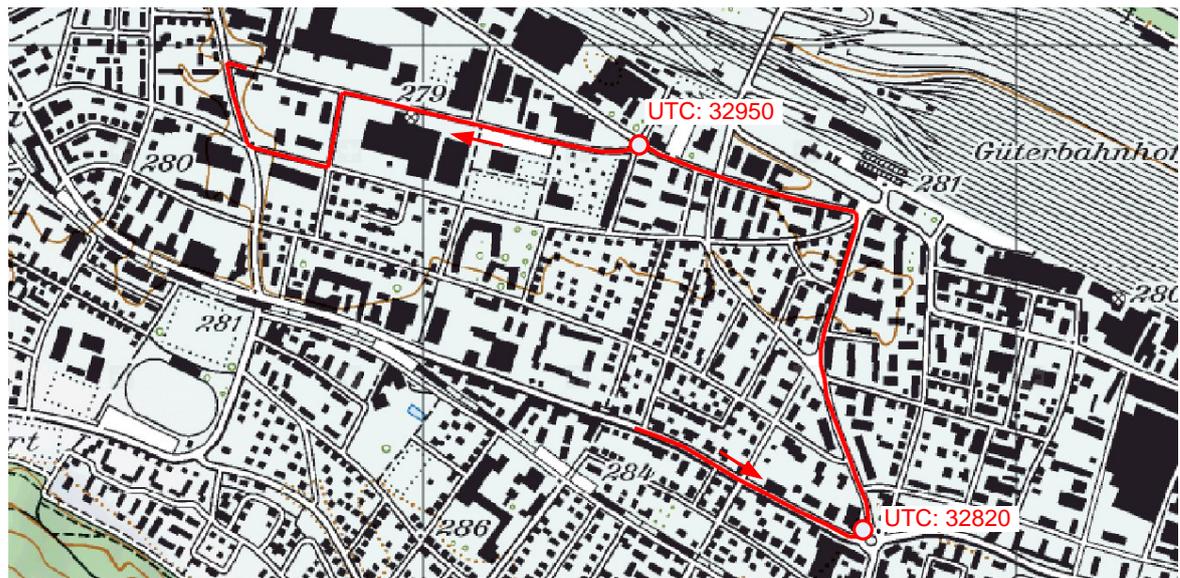


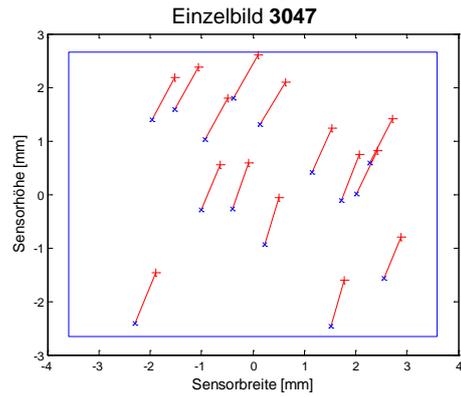
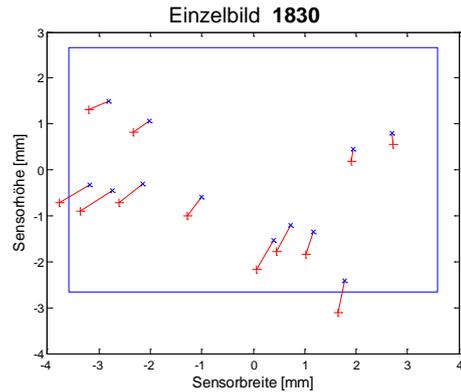
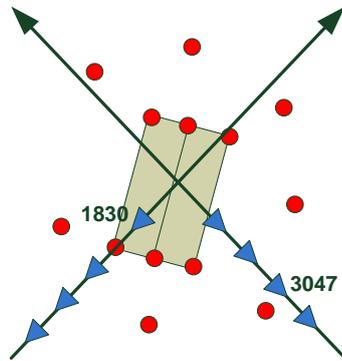
Abbildung 3-5: Übersichtsplan analysierte Trajektorie (oben), Fehlerverhalten der Navigationslösung (AP04) geringer Qualität (unten)

Die detaillierte Analyse der fusionierten Navigationslösung zeigt einen normalverteilten Fehleranteil (Rauschen) sowie einen im hohen Masse zeitlich korrelierten und systematischen Anteil (Signal) (vgl. Abbildung 3-5). Dieses Verhalten lässt sich durch unterschiedliche Effekte erklären. Fehlerquellen sind beispielsweise nicht berücksichtigte systematische Fehlereinflüsse in den GNSS-Pseudorangeabständen oder die zeitlich veränderliche Satellitenkonstellation. Im Weiteren verursachen die hohen Driftraten der MEMS-INS fusioniert mit den GNSS-Pseudorangeabständen systematische Ausrichtungsfehler. Abbildung 3-5 verdeutlicht dieses Fehlerverhalten anhand einer mit dem AP04 über 5 Minuten mobil aufgezeichneten Trajektorie. Deutlich sichtbar ist der zeitlich stark korrelierte absolute Positions- und Gierfehler der Navigationslösung bestimmt aus Sensoren geringer Qualität. Speziell sichtbar sind die grossen Abweichungen bei Richtungsänderungen in der Gierkomponente. Möglicherweise rühren diese Effekte von dem während der Untersuchung nicht berücksichtigtem Steuerinput in die Datenfusion im Fluglagekontroller.

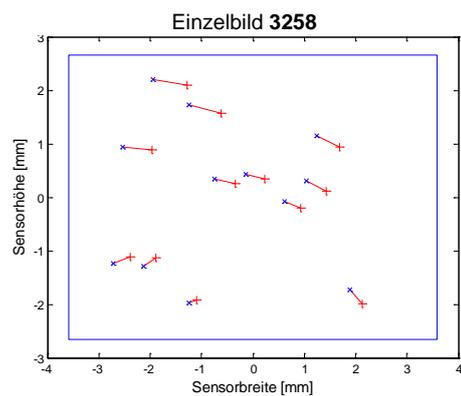
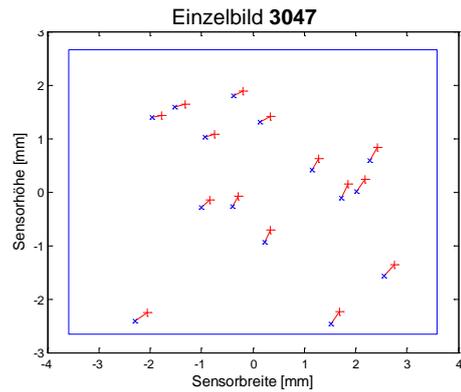
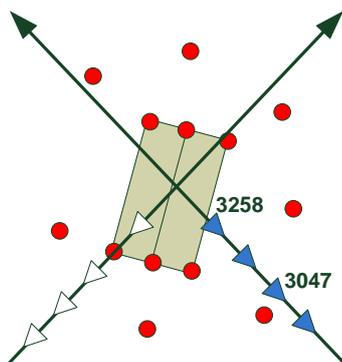
3.2.2.2 Einfluss auf Systemkalibrierung

Das im Abschnitt 3.2.2.1 nachgewiesene systematische Fehlerverhalten der Navigationslösung bestimmt aus MEMS-INS/GNSS-Sensoren geringer Qualität hat grossen Einfluss auf die in Abschnitt 2.3.2.1 eingeführte Systemkalibrierung. Für die Kalibrierung wird vielfach ein Kreuzflug über einem Passpunktfeld durchgeführt. Aufgrund der resultierenden Differenzen zwischen direkt gemessenen und indirekt geschätzten äusseren Orientierungen können die gesuchten Fehlauseinandersetzungen und gegebenenfalls die Hebelarmkomponenten bestimmt werden. Im Falle von INS/GNSS Sensoren hoher Qualität können wir die Fehlauseinschätzung zuverlässig schätzen, da der zufällige Fehler der direkten Georeferenzierungslösung signifikant kleiner ist als die enthaltene Fehlauseinschätzung. Damit können mit einer solchen Sensorkonfiguration die gesuchten Systemkalibrierungsparameter signifikant bestimmt und entsprechend in der Datenverarbeitung berücksichtigt werden, wie dies z.B. die Untersuchungen von Cramer (2001b) zeigen. Wenn wir jedoch denselben Ansatz zur Kalibrierung von Systemen mit Sensoren geringer Qualität anwenden, ist die gesuchte Fehlauseinschätzung mit den enthaltenen systematischen Fehlern der Navigationslösung überlagert. Zusätzlich erschwert wird die Schätzung durch die hohe Korrelation zwischen den verbleibenden systematischen Fehlern der Navigationslösung und der unbekanntem Fehlauseinschätzung und den Hebelarmkomponenten. Dieser Effekt lässt sich mit den beiden Kalibrierungskonfigurationen Abbildung 3-6 illustrieren. In Konfiguration a) werden für die Fehlauseinschätzung Einzelbilder aus beiden Flugstreifen verwendet. Wie das verbleibende Restfehlerbild verdeutlicht, ist es wegen der verbleibenden systematischen Fehler in der Navigationslösung nicht möglich, signifikante Fehlauseinschätzungswinkel zu schätzen. Im Gegensatz dazu werden in Konfiguration b) lediglich kurz aufeinanderfolgende Einzelbilder aus einem Flugstreifen für die Kalibrierung verwendet.

- (a) Acht Einzelbilder aus beiden Flugstreifen werden für die Fehlausrichtungsschätzung verwendet



- (b) Vier Einzelbilder aus einem Flugstreifen werden für die Fehlausrichtungsschätzung verwendet



- Kreuzflugmuster**
- ➔ Fluglinien / Flugrichtung
 - Verwendete Einzelbilder
 - Passpunkte

- Restfehlerbild nach der Kalibrierung**
- x Bildbeobachtungen Passpunkte
 - + Rückprojizierte Passpunkte
 - Restfehler in der Bildebene

Abbildung 3-6: Verbleibende Fehler nach der Systemkalibrierung für zwei Kalibrierungskonfigurationen

In diesem Fall kompensieren, wie das Restfehlerbild zeigt, die geschätzten Fehlausrichtungswinkel den aktuellen systematischen Fehler der Navigationslösung. Damit wird klar ersichtlich, dass mit dem etablierten Systemkalibrierungsverfahren lediglich die aktuelle Fehlausrichtung über eine kurze Zeitperiode bestimmt werden kann und Gültigkeit hat. Die mit den Konfigurationen a) und b) aufgezeigte Verhaltensweise weist klar auf enthaltene zeitlich korrelierte systematische Fehler in der Navigationslösung hin, welche die gesuchte Fehlausrichtung überlagern und mit diesen stark korrelieren. Die Korrelation zwischen Fehler in der äusseren Orientierung und den Fehlausrichtungswinkeln und Hebelarmkomponenten ist mit Hilfe von Wirkfiguren in Ries (2004) detailliert behandelt. Das Fehlerverhalten der Navigationslösung aus Sensoren geringer Qualität verhindert damit eine sinnvolle Bestimmung der Systemkalibrierung mit den eingeführten etablierten Verfahren. Die Behandlung dieser Erkenntnisse verlangt damit nach neuen Verfahren und Methoden zur adäquaten Steigerung der Georegistrierungsgenauigkeit und Kompensation systematischer Fehlereinflüsse von direkt georeferenzierten Videosequenzen, die mit Navigationssensoren geringer Qualität erfasst werden. Ein entsprechender Lösungsansatz wird im nachfolgenden Abschnitt eingeführt.

3.2.2.3 Kontinuierliche Systemkalibrierung

Die erwähnte hohe Korrelation zwischen dem systematischen Fehlerverhalten der Navigationslösung und der unbekanntem Fehlausrichtung verunmöglicht einerseits eine zuverlässige Schätzung dieser Parameter. Andererseits ist die Kompensation des systematischen Fehlausrichtungseinflusses an der Navigationslösung für die Steigerung der Genauigkeit aufgrund des aufgezeigten Fehlerverhaltens der Navigationslösung nicht sinnvoll. Aus diesem Grund werden in den nachfolgenden Betrachtungen die beiden Fehleranteile zu einem Fehlerterm zusammengefasst. Ziel dieses Ansatzes ist es nun, den verbleibenden zeitlich hoch korrelierten systematischen Fehlerterm der Navigationslösung mit Hilfe einer von den verwendeten Navigationssensoren unabhängigen Methode kontinuierlich zu schätzen. Die Lösung der direkten Georeferenzierung kann anschliessend mit dem fortlaufend bestimmten systematischen Fehlerterm korrigiert und verbessert werden. Die konkrete Umsetzung dieses Lösungsansatzes wird im Detail in der zweiten Registrierungsstufe in Abschnitt 3.3.4 behandelt und vorgestellt.

3.3 Integrierte Video-Georeferenzierung - Registrierungsstufe II

Die zweite Registrierungsstufe versucht, die vorliegende Lösung aus der direkten Georeferenzierung mit Hilfe eines integrierten Ansatzes (vgl. Abschnitt 3.1.3) zu verbessern, um eine möglichst optimale Georegistrierungsgenauigkeit der erfassten Videosequenz zu realisieren. Das nachfolgend im Detail eingeführte integrierte Lösungskonzept berücksichtigt die in Abschnitt 3.2.2.2 dargelegten Erkenntnisse und basiert auf der in Abschnitt 3.2.2.3 vorgeschlagenen kontinuierlichen Systemkalibrierung. Nachfolgend wird zuerst die Architektur der zweiten Registrierungsstufe vorgestellt. Anschliessend werden die Teilkomponenten 'Bild-zu-Modell-Zuordnung', 'Bildorientierung und Online-Triangulation' und 'kontinuierliche Fehlerzustandschätzung' eingeführt. Insbesondere bieten die besprochenen Ansätze und Konzepte im Abschnitt 'Bild-zu-Modell-Zuordnung' in Kombination mit der unabhängigen Bildorientierung interessante Lösungsansätze, um die weiteren in der Einleitung erwähnten Anwendungsszenarien zu realisieren.

3.3.1 Architektur des integrierten Lösungsansatzes

Grundidee bildet die Beobachtung der zur Verfügung stehenden Navigationslösung mit Hilfe einer von der INS/GNSS-Sensorik unabhängigen Messmethode. Mit Hilfe der unabhängig durchgeführten Beobachtungen können verbleibende systematische Fehler in der Navigationslösung geschätzt und entsprechend berücksichtigt werden. Die verbleibenden systematischen Fehler der Navigationslösung aus der direkten Georeferenzierung werden über den Kamerapfad variabel, aber mit einer hohen zeitabhängigen Korrelation angenommen. Dieses Verhalten des systematischen Fehleranteils wird mit den in Abschnitt 3.2.2.1 aufgezeigten Erkenntnissen begründet. Als unabhängiger Beobachtungssensor kommt die auf der Aufnahmeplattform verfügbare Mono-Videokamerakonfiguration zur Anwendung. Mit Hilfe indirekter Bildorientierungsverfahren lässt sich die äussere Orientierung aus Einzelbildern der aufgenommenen Videosequenz unabhängig bestimmen. Nachfolgende Abbildung 3-7 zeigt die Architektur und die Teilkomponenten des vorgeschlagenen integrierten Lösungsansatzes.

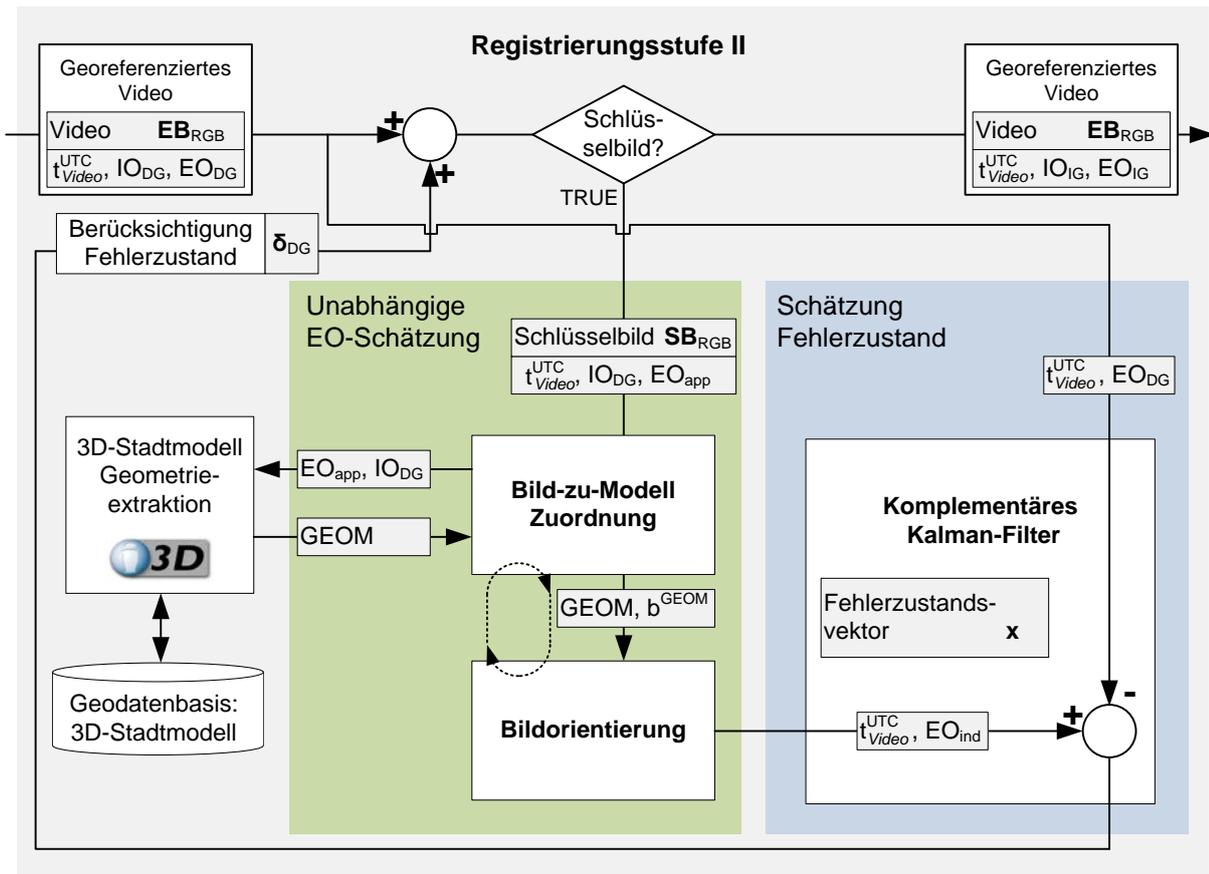


Abbildung 3-7: Architektur Registrierungsstufe II - Integrierte Georeferenzierung

Ausgangslage ist die georeferenzierte Videosequenz als Resultat aus der vorgängig durchgeführten direkten Georeferenzierung, womit für jedes Einzelbild der Sequenz die Zeitepoche bezüglich UTC und das Sensormodell bestehend aus innerer und äusserer Orientierung vorliegen. Ziel ist nun, auf Basis der zur Verfügung stehenden Videosequenz und der vorliegenden 3D-Stadtmodelldaten eine unabhängige Bildorientierung zu bestimmen, mit welcher die gesuchten systematischen Fehler der direkten Georeferenzierung beobachtet werden können. Die über den virtuellen Globus verfügbaren Referenz-Stadtmodelldaten liegen im globalen geodätischen Referenzsystem WGS 84 vor. Liegt am Eingang ein entsprechendes Schlüsselbild SB_{RGB} der Sequenz an, wird folgender Ansatz zur Aktualisierung des aktuellen systematischen Fehlerzustandes ausgeführt. Die erste Teilkomponente bestimmt basierend auf einer Bild-zu-Modell-Zuordnung und in Kombination mit einem photogrammetrischen Bildorientierungsverfahren die gesuchte unabhängige äussere Orientierung EO_{ind} für das Schlüsselbild. Die dafür benötigten Referenz-3D-Vektordaten $GEOM$ können mit Hilfe des virtuellen Globus aufgrund der zur Verfügung stehenden Näherungslösung EO_{app} sehr effizient aus einem bestehenden 3D-Stadtmodell extrahiert werden. Das automatische Bild-zu-Modell-Zuordnungsverfahren sucht die optimale Abbildung zwischen Schlüsselbild und den extrahierten Referenz-3D-Vektordaten. Die Abbildung zwischen extrahierten Bildbeobachtungen b^{GEOM} und Referenzgeometrie $GEOM$ bildet nun die Grundlage für die auf der Perspektiven-Abbildung basierenden unabhängigen Bildorientierung. Die

zweite Teilkomponente schätzt nun aufgrund der zur Verfügung stehenden unabhängig bestimmten \mathbf{EO}_{ind} und der direkt gemessenen äusseren Orientierung \mathbf{EO}_{DG} des Schlüsselbildes den gesuchten aktualisierten systematischen Fehlerterm δ_{DG} . Die systematische Fehlertermschätzung verwendet ein zeitdiskretes Kalman-Filter (vgl. Abschnitt 2.4), welches im Zustandsvektor \mathbf{x} das aktuelle systematische Fehlerverhalten der einzelnen Parameter der äusseren Orientierung modelliert. Der geschätzte Fehlerzustand δ_{DG} kann nun auf die fortlaufenden am Eingang anliegenden Einzelbilder prädiert und für die Video-Georeferenzierung berücksichtigt werden. Daraus resultiert die verbesserte integrierte Georeferenzierungslösung \mathbf{EO}_{IG} für jedes Einzelbild der Sequenz. Für das nächste Schlüsselbild in der Verarbeitung wird nun die verbesserte äussere Orientierung \mathbf{EO}_{IG} für die Bild-zu-Modell-Zuordnung als Näherungslösung \mathbf{EO}_{app} eingeführt, wodurch der Zuordnungsprozess bedeutend vereinfacht wird. Im Unterschied zu den integrierten Ansätzen die in der Photogrammetrie vorgeschlagen werden, wird der verbleibende systematische Fehleranteil hier nicht durch eine lineare Funktion, beispielsweise modelliert durch einen Shift und Drift (Haala, 2005), sondern mittels zeitlich veränderlicher Zufallsbewegung (engl. *Random Walk*) beschrieben. Die Verwendung eines *Random Walk* erlaubt die Beschreibung sämtlicher miteinander hoch korrelierter systematischer Fehleranteile in einem Fehlerterm. Dank dieser Formulierung muss das Verhalten der einzelnen enthaltenen Fehlereinflüsse weder genau bekannt, noch muss eine Funktion gefunden werden, mit welcher diese Einflüsse korrekt beschreibbar sind. Diese Modellierung berücksichtigt damit das in Abschnitt 3.2.2.1 nachgewiesene Fehlerverhalten einer direkt gemessenen Flugtrajektorie mit Navigationssensoren geringer Qualität.

3.3.2 Bild-zu-Modell-Zuordnung

Ziel dieses Verarbeitungsschritts ist die automatische Zuordnung extrahierter Objekte aus einem Einzelbild (Farb- oder Grauwertbild) auf homologe Geometrielemente eines bestehenden 3D-Landschafts- oder Stadtmodells. Das Resultat der Zuordnung definiert für die anschliessende photogrammetrische Orientierungsbestimmung die Bildbeobachtungen auf die zur Verfügung stehende Referenzgeometrie. Die Orientierungsbestimmung kann unabhängig sequentiell nach der automatischen Bild-zu-Modell-Zuordnung erfolgen. Weiter denkbar ist die direkte Integration der Orientierungsbestimmung in das Zuordnungsverfahren. In diesem Fall kann die für die Bildorientierung gesuchte perspektivische Transformation zwischen Bild und Referenzgeometrie direkt als Abbildungsfunktion h verwendet werden.

Ausgangslage für die Bild-zu-Modell-Zuordnung ist eine Näherungslösung der Bildorientierung bezüglich der vorliegenden Geometrielemente im 3D-Stadtmodell. Die Näherungslösung besteht aus den Parametern der äusseren Orientierung, welche zusammen mit der als bekannt vorausgesetzten inneren Orientierung näherungsweise das gesuchte Sensormodell des Einzelbildes definieren. Mit Hilfe einer räumlichen Sichtkegelabfrage, aufgrund des näherungsweise bekannten Sensormodells, können über die Schnittstelle des virtuellen Globus effizient bestehende Geometrielemente aus

einem beliebig grossen digitalen Stadtmodell für den Zuordnungsprozess extrahiert werden. Die extrahierte Referenzgeometrie liegt in der Randdarstellungsrepräsentation (*B-Rep* - engl. *Boundary Representation*) (Zanini, 1998) dreidimensional bezüglich eines globalen geodätischen Referenzsystems vor (z.B. WGS 84). Die Resultatmenge der Sichtkegelabfrage kann zusätzlich durch im Stadtmodell verfügbare Objekt- und Elementsachdaten eingeschränkt werden, was den Zuordnungsprozess in bestimmten Anwendungen erheblich vereinfachen kann.

Nachfolgender Abschnitt vermittelt als erstes einen Überblick über mögliche Zuordnungsmethoden, welche für die Lösung der Problemstellung in Betracht gezogen werden können. Der zweite Abschnitt beschreibt im Anschluss im Detail die theoretischen Grundlagen des gewählten und implementierten Zuordnungsverfahrens.

3.3.2.1 Lösungsansätze

Für die Lösung des Zuordnungsproblems können unterschiedliche Verfahren in Betracht gezogen werden. Das optimale Verfahren hängt sehr stark von der Qualität der zur Verfügung stehenden Ausgangsdaten ab. Für die Lösung der Aufgabenstellung ist ein automatisches Verfahren gesucht, welches die Zuordnung zwischen einer Bildaufnahme von von Menschenhand erbauter Objekte und dem digitalen 3D-Abbild in Form eines Vektormodells dieser Objekte bestimmt. Im Kontext der im Abschnitt 2.2.1 eingeführten Klassifizierung von Zuordnungsverfahren kann das Problem mit der Zuordnungsebene Bild-zu-Modell beschrieben werden. Wie Tabelle 2-2 verdeutlicht, können unterschiedliche Zuordnungsmethoden für deren Lösung genutzt werden. Nachfolgend werden drei konkrete Ansätze unterschiedlicher Zuordnungsmethoden diskutiert und hinsichtlich der Verwendbarkeit für die Lösung der vorliegenden Problemstellung beurteilt.

Zuordnung von Kanten - Merkmalsbasierte Zuordnung (FBM)

Dieser Ansatz sucht eine optimale Abbildung h , welche Kanten-Primitiven zweier Modellbeschreibungen einander zuweist. Das vorliegende Bild-zu-Modell-Zuordnungsproblem wird hier auf die in Gleichung (4) dargestellte vereinfachte Form zurückgeführt. Nachfolgende Abbildung 3-8 zeigt den untersuchten Zuordnungsprozess: als erstes wird die 3D-Modellgeometrie mit dem näherungsweise vorliegenden Sensormodell bestehend aus \mathbf{EO}_{app} und \mathbf{IO}_{DG} in die Bildebene der Kamera abgebildet, aus welcher anschliessend die Modellbeschreibung \mathbf{m}^{mo} abgebildet werden kann. Parallel dazu werden mit Hilfe von *midlevel* Bildverarbeitungsalgorithmen Kanten aus dem vorliegenden Schlüsselbild extrahiert. Dazu werden mit einem Gradienten-Filter (z.B. (Canny, 1983)) als erstes Kantenpixel extrahiert. Ausgehend vom Resultat dieser Vorfilterung werden mit Hilfe der linearen Hough-Transformation (HT) (Duda und Hart, 1972) Kanten in Form von 2D-Vektoren extrahiert, aus denen die Modellbeschreibung \mathbf{m}^{img} abgeleitet wird.

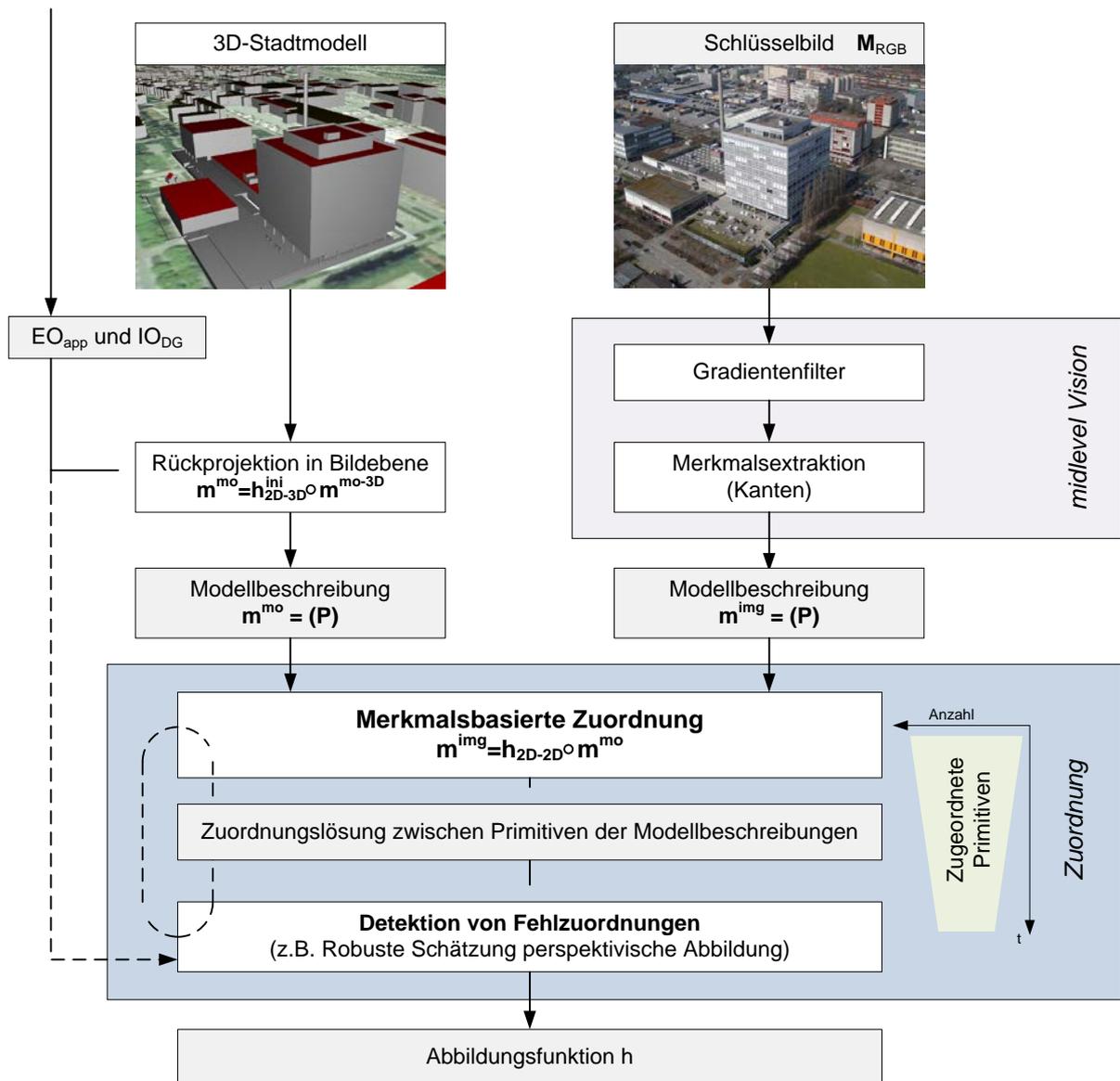


Abbildung 3-8: Bild-zu-Modell-Zuordnung mit markmalsbasiertem Ansatz

Die Zuordnung erfolgt anschliessend basierend auf den zur Verfügung stehenden Merkmalsdeskriptoren der extrahierten Kanten-Primitiven nach dem in Abschnitt 2.2.2.1 eingeführten Verfahren von McIntosh und Mutch (1988). Da der Merkmalsvektor der Kanten-Primitiven der Modellbeschreibung m^{mo} lediglich durch geometrische Eigenschaften gebildet werden kann, können die vorgeschlagenen radiometrischen Eigenschaften in der Nachbarschaft der extrahierten Kante als Zuordnungskriterium nicht verwendet werden. Der Zuordnungsdeskriptor einer Primitive p_i^k von m^{img} und m^{mo} besteht aus den Eigenschaften Kanten-Länge, Kanten-Orientierung, Lage des Kanten-Mittelpunktes und wird mit der Anzahl paralleler Kanten links und rechts der jeweiligen Primitive erweitert. Anhand der zusätzlich eingeführten Eigenschaft kann auf einfache Weise die Nachbarschaft einer Kante in der Zuordnungsbewertung berücksichtigt werden. Die mit diesem Verfahren erzielbare Qualität der Zuordnungslösung ist im Wesentlichen abhängig von der Genauigkeit des für die Rückprojektion benötigten initialen Sensormodells. Damit die Zuordnungslösung für die

weitere Datenverarbeitung genutzt werden kann, müssen Fehlzuweisungen detektiert und automatisch aus der Zuordnungslösung eliminiert werden. Zu diesem Zweck wird ein robustes Schätzverfahren vorgeschlagen, welches direkt die perspektivischen Abbildungsparameter des Schlüsselbildes aufgrund der Zuordnungslösung bestimmt. Wird davon ausgegangen, dass die Zuordnung einen hohen Anteil Fehlzuweisungen enthält, ist ein robustes Schätzverfahren mit hohem Bruchpunkt angezeigt. Ein geeignetes Verfahren für diese Ausgangslage ist beispielsweise der von Fischler und Bolles (1981) vorgeschlagene Ansatz *Random Sample Consensus* (RANSAC). Die robust geschätzten perspektivischen Abbildungsparameter h des Schlüsselbildes entsprechen gerade den für die integrierte Video-Georegistrierung gesuchten unabhängig bestimmten äusseren Orientierungsparameter \mathbf{EO}_{ind} .

Zuordnung topologischer Modelle - Relationale Zuordnung (RM)

Im Gegensatz zum oben eingeführten rein merkmalsbasierten Lösungsansatzes nutzt dieser Ansatz zur Lösung des Zuordnungsproblems zwei topologische Modellbeschreibungen. In der Bewertung für die Suche der optimalen Zuordnung wird nicht nur ausschliesslich der Merkmalsdeskriptor einer Modellprimitive, sondern es werden auch modellierte Relationen zwischen den Primitiven berücksichtigt. Abbildung 3-9 zeigt das Konzept des vorgeschlagenen Ansatzes, welcher die Bild-zu-Modell-Zuordnungsaufgabe nach der in Gleichung (4) formulierten Betrachtungsweise löst. Zu Beginn der Verarbeitung wird die 3D-Modellgeometrie mit Hilfe des näherungsweise vorliegenden Sensormodells bestehend aus \mathbf{EO}_{app} und \mathbf{IO}_{DG} in die Bildebene der Kamera abgebildet. Die im 3D-Stadtmodell modellierte Topologie kann direkt genutzt werden, um die gesuchte topologische Modellbeschreibung \mathbf{m}^{mo} zu bilden. Parallel dazu wird ein *high-level* Bildverarbeitungsprozess benötigt, welcher aus dem vorliegenden Schlüsselbild Kanten-Primitiven und die darin enthaltene Topologie extrahiert. Mit Hilfe dieser Information lässt sich die für den Zuordnungsprozess gesuchte Modellbeschreibung \mathbf{m}^{img} ableiten. Die Zuordnung erfolgt mit Hilfe eines modifizierten Ansatzes nach der in Abschnitt 2.2.2.3 eingeführten relationalen Zuordnung nach Vosselman (1992). Eine Modellbeschreibung $m^k = (P, R)$ besteht aus Kanten-Primitiven und Beziehungen zwischen den einzelnen Primitiven. Als Deskriptoren für die Kanten werden Länge, Orientierung sowie die Lage des Kantenmittelpunktes verwendet. Die Topologie wird durch die beiden Beziehungstypen beschrieben: sind zwei Kanten miteinander verbunden und dem Abstand zwischen zwei parallelen Primitiven. Die optimale Abbildungsfunktion h zwischen den beiden Modellbeschreibungen, welche die Primitiven der Modellbeschreibung \mathbf{m}^{img} den Primitiven der Modellbeschreibung \mathbf{m}^{mo} zuordnet, wird mit Hilfe einer informierten Baumsuche gelöst. Dieser Algorithmus bestimmt im Gegensatz zur merkmalsbasierten Lösungsstrategie zu Beginn wenige sichere Zuordnungen und baut die Zuordnungslösung im Verlaufe der Optimierung laufend aus.

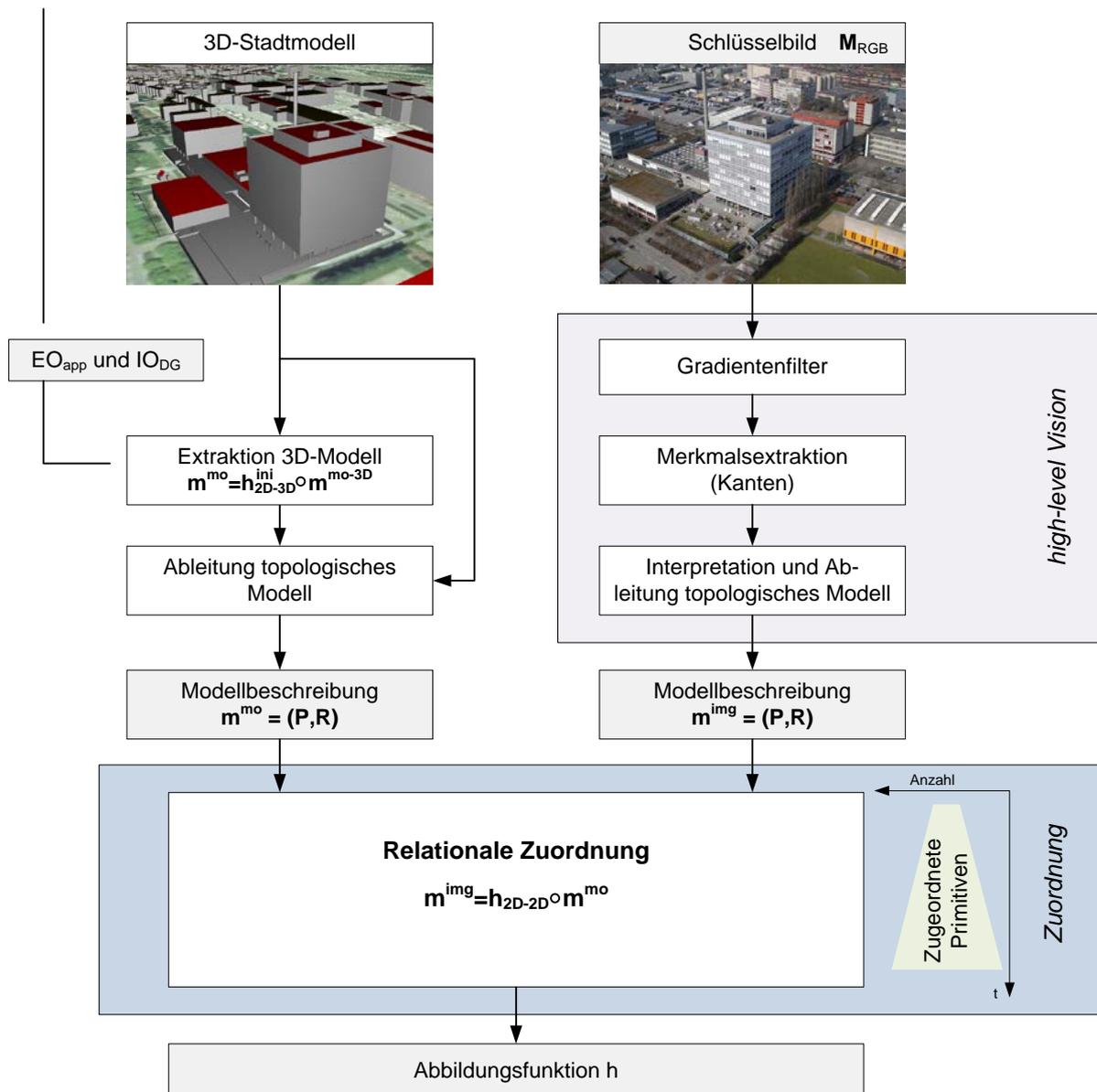


Abbildung 3-9: Bild-zu-Modell-Zuordnung - Relationaler Ansatz

Anpassung der EO-Parameter - Energiefunktional Optimierung (EFO)

Dieser Ansatz schätzt im Zuordnungsprozess direkt die gesuchten äusseren Orientierungsparameter EO_{ind} und verwendet als Abbildungsfunktion h die Kollinearitätsbeziehung zwischen Kanten des 3D-Modells und dessen Abbildung im Schlüsselbild im Sinne von Gleichung (3). Konzeptionell ist dieser Ansatz in Abbildung 3-10 dargestellt. Der hier vorgeschlagene Ansatz basiert auf dem in Abschnitt 2.2.2.2 eingeführten Chamfer-Algorithmus (Borgefors, 1988), welcher zu den Energiefunktional optimierenden Zuordnungsmethoden zählt. Der von Borgefors vorgeschlagene Algorithmus verwendet aufgrund der verwendeten Energiefunktion für die Bestimmung der äusseren Orientierungsparameter ein direktes Optimierungsverfahren, welches im Allgemeinen eine geringe Konvergenzgeschwindigkeit aufweist. Wie die Abbildung 3-10 zeigt, wird bei dem hier vorgeschlagenen Ansatz das von Borgefors verwendete

prepolygon Bild (vgl. Abbildung E-1) durch das ebenfalls distanztransformierte Bild $g^0(x,y)$ ersetzt. Damit wird erreicht, dass mit $g^0(x,y)$ eine stetige Grauwertfunktion der Modellbeschreibung m^{mo} vorliegt, die es erlaubt, die gesuchten äusseren Orientierungsparameter der Abbildungsfunktion h mit Hilfe eines lokalen Gradientenverfahrens zu optimieren. Das Problem kann unter dieser Voraussetzung als optimale Anpassung der Distanz-Grauwertfunktionen im Sinne der kleinsten Distanzdifferenz $e(x,y)$ - Quadratesumme aufgefasst werden.

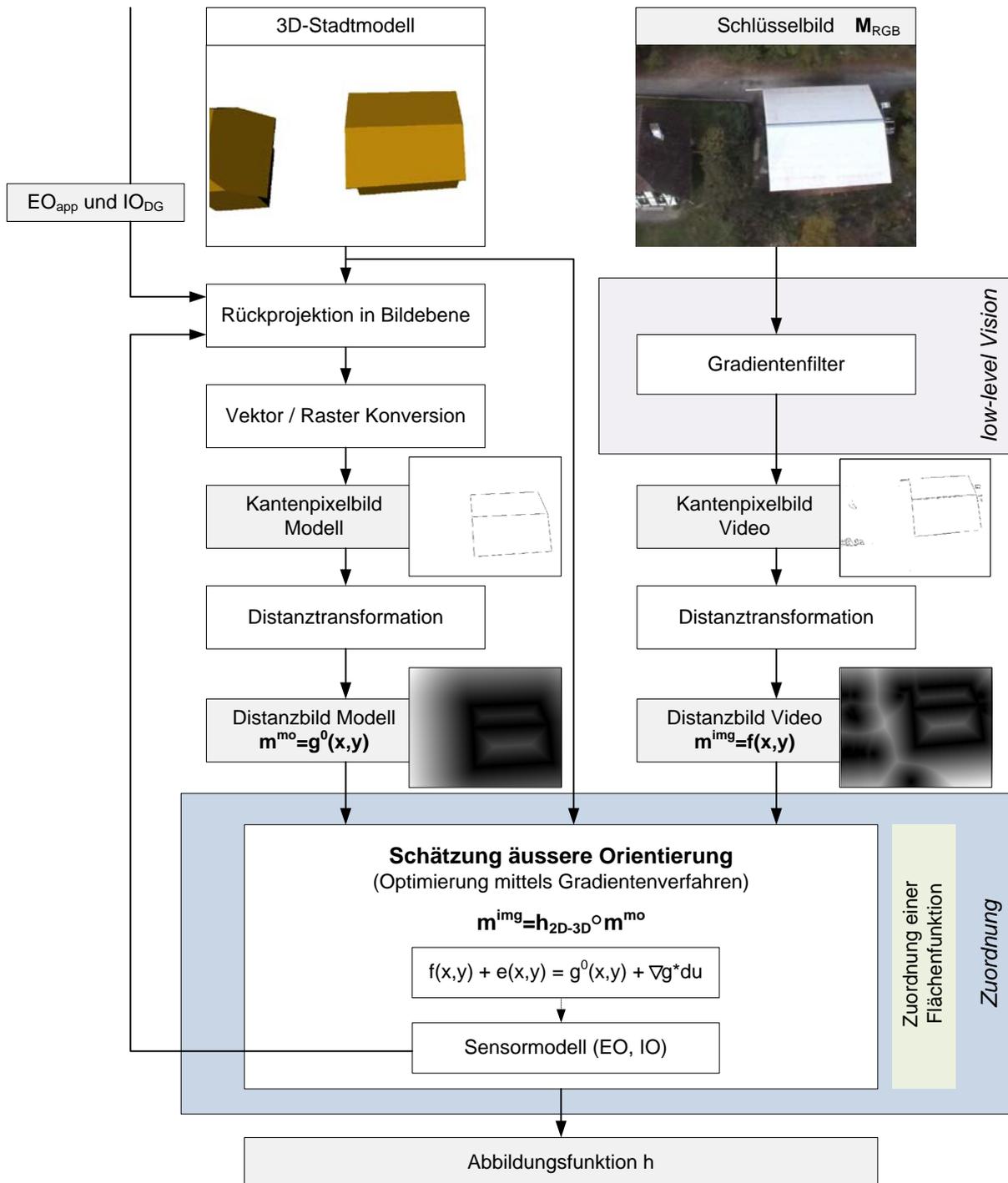


Abbildung 3-10: Bild-zu-Modell-Zuordnung mittels Energiefunktionaloptimierung

Die dargestellte Ausformulierung der Problemstellung ist an die kleinste Quadrate basierte Zuordnung (ABM) angelehnt, und betrachtet im Gegensatz dazu die Problemstellung global über das gesamte Bild. Gesucht wird demzufolge ein Distanzbild $g(x,y)$, das bis auf ein zufälliges Rauschen $e(x,y)$ mit dem Distanzbild $f(x,y)$ der Modellbeschreibung \mathbf{m}^{img} identisch ist.

$$f(x, y) - e(x, y) = g(x, y) \quad (25)$$

Jeder Distanzwert $g_i(x,y)$ der Modellbeschreibung \mathbf{m}^{mo} hängt von den unbekanntem Parametern der äusseren Orientierung $u = (X_0, Y_0, Z_0, \omega, \varphi, \kappa)^T$ wie folgt ab:

$$g_i(x, y) = \sqrt{(x - x_e)^2 + (y - y_e)^2} \quad (26)$$

mit

$$\begin{pmatrix} x_e \\ y_e \\ -c \end{pmatrix} = \frac{1}{\lambda} \cdot \mathbf{R}_{\text{global}}^{\text{img}}(\omega, \varphi, \kappa) \cdot \left[\mathbf{E}_{\text{global}} - \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} \right]$$

Darin beschreibt x_e und y_e die Lage des nächst gelegenen Kantenpixels einer ins Bild zurückprojizierten Modellkante E_{global} . Für die Optimierung der gesuchten äusseren Orientierungsparameter wird das Gauss-Newton Verfahren verwendet (vgl. Abschnitt 2.3.1.1), welches im Allgemeinen eine schnelle Konvergenz verspricht. Die dafür benötigte Beobachtungsgleichung Gleichung (26) ist bezüglich der unbekanntem Parameter u nicht linear und ist zu linearisieren. Daraus folgen die linearisierten Verbesserungsgleichungen:

$$f(x, y) - e(x, y) = g^0(x, y) + \sum_{i=1}^u \left(\left(\frac{\partial g^0(x, y)}{\partial x_e} \cdot \frac{\partial x_e}{\partial u_i} + \frac{\partial g^0(x, y)}{\partial y_e} \cdot \frac{\partial y_e}{\partial u_i} \right) \cdot du_i \right) \quad (27)$$

Mit jedem Iterationsschritt werden aufgrund der geschätzten unbekanntem Zuschläge du die initialen äusseren Orientierungsparameter \mathbf{EO}_{app} neu bestimmt und eine neue Modellbeschreibung \mathbf{m}^{mo} abgeleitet. Erfahrungen mit diesem Ansatz haben gezeigt, dass das Verfahren speziell in jenen Fällen nicht robust ist, in denen im Schlüsselbild viele Kantenpixel detektiert werden, die nicht einer im Modell vorhandenen Kante zugeordnet werden können. Vielfach konvergiert in solchen Situationen das Verfahren in einem beliebigen lokalen Minimum. Der Ansatz kann aber durch eine robuste Erweiterung beispielsweise durch die Implementierung eines M-Schätzers wie beispielsweise in Wicki (1998) verwendet oder durch eine bufferbasierte Vorfilterung der detektierten Kantenpixel stabilisiert werden. Der Berechnungsaufwand lässt sich zusätzlich unter Verwendung einer Bildpyramide im Iterationsprozess reduzieren.

Beurteilung und Schlussfolgerung

Zusammenfassend können die drei vorgestellten Lösungsansätze nach unterschiedlichen Kriterien analysiert und beurteilt werden. Untersuchungen zeigten, dass hauptsächlich die Genauigkeit der vorliegenden initialen äusseren Orientierung \mathbf{EO}_{app} die Lösbarkeit des Zuordnungsproblems bestimmt und damit für die Wahl des Lösungsansatzes entscheidend ist. Nachfolgende Tabelle vergleicht die Ansätze hinsichtlich Ausgangslage, Bildverarbeitungsebene, Zuordnungsstrategie, Rechenkomplexität und Robustheit.

Kriterium	FBM	RM	EFO
Einfluss Güte Näherungslösung	hoch	tief	mittel
Bildverarbeitungsebene	mid-level	high-level	low-level
Zuordnungsstrategie	Ausdünnung zugeordneter Merkmale	Ausbau zugeordneter Merkmale	Funktionaler Zusammenhang
Optimierungs-Algorithmus	-	Informierte Baumsuche	Gradientenverfahren
Rechenkomplexität	mittel	hoch	hoch
Robuste Verarbeitung	erforderlich	implizit integriert	erforderlich

Tabelle 3-2: Gegenüberstellung der untersuchten Bild-zu-Modell-Zuordnungsansätze

Bei der typischerweise vorliegenden Genauigkeit der Näherungslösung aus der direkten Georeferenzierung zeigen die Zuordnungsergebnisse der Lösungsansätze FBM und EFO starke Defizite. Einzig der relationale Ansatz, der im Speziellen zusätzliche topologische Information zwischen den Primitiven der Modellbeschreibung berücksichtigt, ermöglicht die erfolgreiche Zuordnung. Wie in Tabelle 3-2 zusammenfassend aufgezeigt, ist das Zuordnungsproblem damit allgemein lösbar, jedoch zu Lasten eines höheren Rechenaufwandes im Vergleich zum merkmalsbasierten Ansatz. Die Ableitung einer topologisch korrekten Modellbeschreibung \mathbf{m}^{img} bedingt im Gegensatz zu den beiden anderen Ansätzen eine vorgängige komplexe Merkmalsextraktion und Interpretation. Die Güte der abgeleiteten Modellbeschreibung beeinflusst im Wesentlichen das Ergebnis des Zuordnungsprozesses. Die beiden Ansätze merkmalsbasierte Zuordnung und erweiterter Chamfer-Algorithmus können bei einer ausreichend genauen initialen Bildorientierung in Kombination mit einer robusten Schätzung erfolgreich für die Problemlösung eingesetzt werden. Aufgrund dieser Erkenntnisse wird für die Anwendung im Zusammenhang mit Navigationssensoren geringer Qualität die relationale Bild-zu-Modell-Zuordnung vorgeschlagen und verwendet. Jedoch könnte nach erfolgreicher Initialisierung des aktuellen Fehlerzustandes der direkt gemessenen äusseren Orientierung das Zuordnungsproblem mit dem weniger aufwändigen und komplexen merkmalsbasierten Ansatz sinnvoll gelöst werden.

3.3.2.2 Relationale Bild-zu-3D-Stadtmodell-Zuordnung

Nachfolgend wird der in Abschnitt 3.3.2.1 eingeführte relationale Zuordnungsansatz im Detail vorgestellt und besprochen. Der vorgeschlagene Zuordnungsansatz kann nicht nur für die integrierte Video-Georegistrierung angewendet werden, sondern er eignet sich allgemein für die Lösung einer Aufgabenstellung, in welcher eine Einzelaufnahme mit näherungsweise bekannter äusserer Orientierung auf bestehende 3D-Stadtmodelldaten registriert werden soll. Mögliche Anwendungen für diesen Algorithmus sind in Abschnitt 1.1 diskutiert. Die Bild-zu-3D-Stadtmodell-Zuordnung kann im Zusammenhang mit der im nachfolgenden Abschnitt diskutierten Online-Triangulation auch als Algorithmus aufgefasst werden, welcher bestehenden 3D-Vektoren automatisch Bildstrukturen zuordnet, aus welchen anschliessend im photogrammetrischen Sinne die absolute Orientierung des Bildblocks bestimmt werden kann (Cho, 1996).

Aufbauend auf dem in Abbildung 3-9 dargelegten Konzept wird nachfolgend die verwendete topologische Modellbeschreibung besprochen. Weiter wird auf die automatische Ableitung der Modellbeschreibungen \mathbf{m}^{mo} und \mathbf{m}^{img} aus den extrahierten 3D-Stadtmodelldaten und der Einzelaufnahme eingegangen. Im Anschluss wird die gewählte Evaluationsfunktion eingeführt, welche den Nutzen einer entsprechenden Abbildungsfunktion h zwischen den Modellbeschreibungen beurteilt. Die Optimierung der Evaluationsfunktion mittels informierter Baumsuche wird am Ende dieses Abschnittes behandelt.

Topologische Modellbeschreibung

Die vorgeschlagene topologische Modellbeschreibung wird aus n Primitiven und l Beziehungen wie folgt gebildet:

$$m = (P, R) \quad (28)$$

$$P = \{p_1, p_2, \dots, p_n\} \quad (29)$$

$$R = \{r_1^t(p_i, p_j), r_2^t(p_i, p_j), \dots, r_l^t(p_i, p_j)\} \quad (30)$$

darin bezeichnen

p_i : Kanten-Primitive

r_i^t : Beziehung vom Typ t zwischen zwei Primitiven der Menge P

Abbildung 3-11 veranschaulicht die in der Bildebene definierte topologische 2D-Modellbeschreibung bestehend aus den eingeführten Primitiven und Beziehungen.

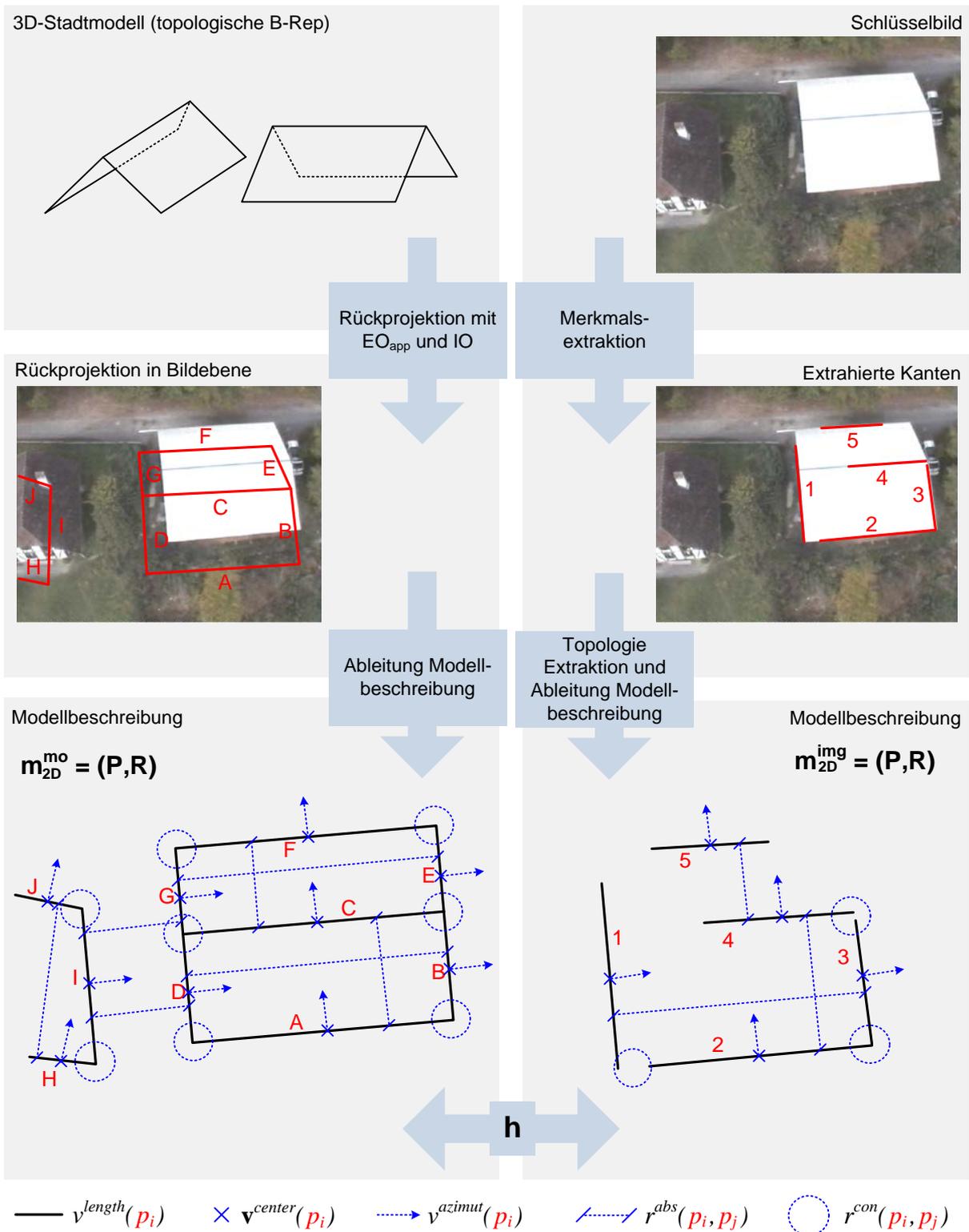


Abbildung 3-11: Extraktion topologische Modellbeschreibung in der Bildebene

Die Modellbeschreibung verwendet Kantenobjekte als Primitiven. Eine Kante kann durch deren Anfangs- und Endpunkt (Stützpunkte) definiert im Bildkoordinatensystem dargestellt werden.

$$E = [e^{start}, e^{end}] \quad \text{mit} \quad e^{start} = \begin{pmatrix} x^{start} \\ y^{start} \end{pmatrix}, \quad e^{end} = \begin{pmatrix} x^{end} \\ y^{end} \end{pmatrix} \quad (31)$$

Ergänzend zu dieser Repräsentation lässt sich die Kante auch mit folgender Geradendarstellung beschreiben.

$$d = x \cdot \sin(\varphi) + y \cdot \cos(\varphi) \quad (32)$$

mit

- d : Abstand der Geraden zum Koordinatenursprung (Bildhauptpunkt)
 φ : Orientierung der Normalen zur Geraden
 x, y : Punkt auf der Geraden

Aus den beiden Darstellungsformen lässt sich folgender Merkmalsvektor einer Kanten-Primitiven

$$p = \begin{pmatrix} \mathbf{v}^{center}(E) \\ \mathbf{v}^{azimut}(E) \\ \mathbf{v}^{length}(E) \end{pmatrix} \quad (33)$$

gebildet aus den Eigenschaften

$$\mathbf{v}^{center}(E) = \begin{pmatrix} x^{mid} \\ y^{mid} \end{pmatrix} = \begin{cases} x^{mid} = (x^{start} + x^{end})/2 \\ y^{mid} = (y^{start} + y^{end})/2 \end{cases} \quad (34)$$

$$\mathbf{v}^{azimut}(E) = \varphi \quad (35)$$

$$\mathbf{v}^{length}(E) = \sqrt{(x^{end} - x^{start})^2 + (y^{end} - y^{start})^2} \quad (36)$$

ableiten. Die zusätzlich benötigte topologische Information zwischen den Kanten-Primitiven wird mit Hilfe der beiden nachfolgenden Beziehungstypen in die Modellbeschreibung integriert.

$$r^{con}(p_i, p_j) = \begin{cases} 1 \\ 0 \end{cases} \quad \text{für} \quad \begin{cases} (e_i^{start} = e_j^{start}) \vee (e_i^{end} = e_j^{end}) \vee (e_i^{start} = e_j^{end}) \vee (e_i^{end} = e_j^{start}) \\ (e_i^{start} \neq e_j^{start}) \wedge (e_i^{end} \neq e_j^{end}) \wedge (e_i^{start} \neq e_j^{end}) \wedge (e_i^{end} \neq e_j^{start}) \end{cases} \quad (37)$$

$$r^{abs}(p_i, p_j) = \begin{cases} |d_i - d_j| \\ \Omega = \{ \} \end{cases} \quad \text{für} \quad \begin{cases} |\varphi_i - \varphi_j| \leq \Delta\Phi \\ |\varphi_i - \varphi_j| > \Delta\Phi \end{cases} \quad (38)$$

Der Beziehungstyp r^{con} definiert ein Kantenpaar welches durch einen gemeinsamen Stützpunkt verbunden ist. Mit Hilfe des zweiten Beziehungstyps r^{abs} kann der Abstand zwischen zwei parallel ausgerichteten Kanten definiert werden. Dabei definiert der Schwellwert $\Delta\Phi$ ein paralleles Kantenpaar. Mit Hilfe dieser beiden Beziehungstypen werden sämtliche Kantenpaare in einer Modellbeschreibung festgelegt, die einen gemeinsamen Stützpunkt aufweisen oder zueinander parallel ausgerichtet sind.

Die beiden Modelle \mathbf{m}^{mo} und \mathbf{m}^{img} müssen wie in Abbildung 3-9 bereits eingeführt nach der vorgestellten topologischen Modellbeschreibung aus dem ausgeschnittenen 3D-Stadtmodell sowie aus dem Schlüsselbild abgeleitet werden. Aus dem als Randdarstellungsrepräsentation vorliegenden 3D-Stadtmodellausschnitt lässt sich \mathbf{m}^{mo}

einfach bilden. Die jeweiligen Stützpunkte einer Kante im Stadtmodellausschnitt können aufgrund des näherungsweise bekannten Sensormodells mit Hilfe der Kollinearitätsgleichung (C-2) in die virtuelle Bildebene der Modellbeschreibung abgebildet werden. Mit zusätzlichem Berechnungsaufwand kann festgestellt werden, welche Kanten von näher liegenden Objektflächen verdeckt und damit für die Modellbeschreibung nicht benötigt werden. Liegt ein semantisches 3D-Stadtmodell als Grundlage bereit, könnte die Modellbeschreibung mit Hilfe der vorliegenden Informationen zu den einzelnen Objektoberflächen weiter eingegrenzt und nur aus Kanten gebildet werden, welche sich für eine automatische Extraktion aus dem Schlüsselbild eignen (z.B. Dachflächen). Aus den abgebildeten Kanten E_i in der virtuellen Bildebene können die gesuchten m Merkmalsvektoren p_i mit den Gleichungen (34) bis (36) gebildet werden. Liegt das Stadtmodell in einer topologischen Randdarstellungsrepräsentation¹ (Mäntylä, 1984) vor, resultieren die enthaltenen o Beziehungen r_i^{con} direkt aus dieser Information. Ist die Topologie nicht explizit modelliert, müssen diese Beziehungen wie nachfolgend für \mathbf{m}^{img} beschrieben abgeleitet werden. Die p Beziehungen r_i^{abs} können mit Hilfe der Gl. (38) gebildet werden. Damit bildet $m^{mo} = \{p_1, p_2, \dots, p_m, r_1^{con}, r_2^{con}, \dots, r_o^{con}, r_1^{abs}, r_2^{abs}, \dots, r_p^{abs}\}$ das gesuchte komplette Modell. Demgegenüber gestaltet sich die automatische Extraktion der Modellbeschreibung \mathbf{m}^{img} aus dem Schlüsselbild komplizierter und benötigt wie in Abbildung 3-9 dargestellt einen *high-level* Bildverarbeitungsprozess. Im Speziellen dieser Teil ist für den anschliessenden erfolgreichen Zuordnungsprozess von grosser Bedeutung. Die dafür vorgeschlagene Vorgehensweise verwendet Standardalgorithmen aus der Bildverarbeitung und wird in dieser Arbeit nicht weiter speziell untersucht und optimiert. In einem ersten Schritt werden mittels Gradientenfilter Kantenpixel aus dem Schlüsselbild extrahiert. Dazu wird Wahlweise der Sobel-Filter (Jähne, 2005, Kap. 12.7.3) oder der Canny Kanten-Filter (Canny, 1983) eingesetzt. Anschliessend werden mit Hilfe einer geradenbasierten HT (Duda und Hart, 1972) die enthaltenen Kanten E_i im Bild bestimmt. Aus dieser Repräsentation werden nun die gesuchten n Kanten-Primitiven p_i gebildet. Damit die im Schlüsselbild enthaltenen q Beziehungen r_i^{con} bestimmt werden können, müssen vorgängig gemeinsame Stützpunkte der Kanten-Primitiven p_i identifiziert werden. Dazu wird ein Algorithmus eingesetzt der eine verbundene Stützpunkt- und Kantenlistenstruktur (Mäntylä, 1984) aufbaut, welche die gesuchte topologische Information implizit speichert. Dieser Algorithmus verwendet für die Identifikation gemeinsamer Kantenstützpunkte einen Schwellwert, welcher den zulässigen maximalen euklidischen Abstand zweier Kantenstützpunkte festlegt. Dieser Schwellwert muss empirisch festgelegt werden und hängt stark von der aufgenommenen Szene und der Videodatenqualität ab. Untersuchungen haben gezeigt, dass mit einem Schwellwert von 1/15 der Schlüsselbildbreite im Allgemeinen ein zufriedenstellendes Resultat erreicht wird. Bei der Verwendung dieses Algorithmus für die Extraktion der Beziehungen für die Modellbeschreibung \mathbf{m}^{mo} aus einem nicht topologischen 3D-Stadtmodell, kann dieser Schwellwert kleiner gewählt werden. Im Anschluss können die zur Vervollständigung des Modells \mathbf{m}^{img} mit Gl. (38) benötigten r Beziehungen r_i^{abs} aus

¹ Stützpunkt-Kanten-Polygon Listendarstellung

den extrahierten Kanten E_i bestimmt werden. Die vollständige Modellbeschreibung lautet damit: $m^{img} = \{p_1, p_2, \dots, p_n, r_1^{con}, r_2^{con}, \dots, r_q^{con}, r_1^{abs}, r_2^{abs}, \dots, r_r^{abs}\}$.

Definition der Evaluationsfunktion

Die Übereinstimmung einer Zuordnung zwischen den beiden Modellbeschreibungen

$$m_{2D}^{img} = h \circ m_{2D}^{mo} \quad (39)$$

wird mit Hilfe des daraus resultierenden Nutzens mit Hilfe der nachfolgend eingeführten Evaluationsfunktion beschrieben. Die in Gl. (39) verwendete Abbildungsfunktion h beschreibt in diesem Kontext die Menge der s einander aus den beiden Modellen zugeordneten Primitiven z_i .

$$h = \{z_1(p_i^{img} \leftrightarrow p_i^{mo}), z_2(p_i^{img} \leftrightarrow p_i^{mo}), \dots, z_s(p_i^{img} \leftrightarrow p_i^{mo})\} \quad (40)$$

Einschränkend werden Mehrfachzuordnungen zwischen den m Primitiven p_i^{mo} und n Primitiven p_i^{img} in h ausgeschlossen. Jedoch werden leere Zuordnungen $z_i(p_i \leftrightarrow \{\})$ (*Wildcards*) für Primitiven einer Modellbeschreibung unterstützt. Ziel ist es nun, die entsprechende Abbildungsfunktion zu finden, welche die Evaluationsfunktion maximiert. Die Nutzenfunktion hat im Gegensatz zur Kostenfunktion den Vorteil, dass negative Nutzen für Fehlzuordnungen berücksichtigt werden können (Vosselman, 1992). Die vorgeschlagene Evaluationsfunktion basiert auf nachfolgend eingeführten Funktionen, die es erlauben, den Nutzen für einander zugeordnete Primitiven z_i und daraus resultierende Beziehungen aus den beiden Modellbeschreibungen zu bestimmen.

$$N^{disp}(p_i^{mo}, p_j^{img}) = \frac{\max^{disp} - |\mathbf{v}^{center}(p_j^{img}) - \mathbf{v}^{center}(p_i^{mo})|}{\max^{disp}} \quad (41)$$

$$N^{azi}(p_i^{mo}, p_j^{img}) = \frac{\max^{\Delta azi} - |v^{azimut}(p_j^{img}) - v^{azimut}(p_i^{mo})|}{\max^{\Delta azi}} \quad (42)$$

$$N^{len}(p_i^{mo}, p_j^{img}) = \begin{cases} 0 & \text{für } v^{length}(p_j^{img}) < \max^{\Delta len} \cdot v^{length}(p_i^{mo}) \\ -1 & \text{sonst} \end{cases} \quad (43)$$

$$N^{con}(r^{con}(p_i^{mo}, p_j^{mo}), r^{con}(p_k^{img}, p_l^{img})) = \begin{cases} 1 & r^{con}(p_i^{mo}, p_j^{mo}) = 1 \wedge r^{con}(p_k^{img}, p_l^{img}) = 1 \\ 0 & \text{für } \text{sonst} \\ -1 & r^{con}(p_i^{mo}, p_j^{mo}) = 0 \wedge r^{con}(p_k^{img}, p_l^{img}) = 1 \end{cases} \quad (44)$$

$$N^{abs}(r^{abs}(p_i^{mo}, p_j^{mo}), r^{abs}(p_k^{img}, p_l^{img})) = \begin{cases} \text{mer}^{abs} & \text{für } r^{abs}(p_i^{mo}, p_j^{mo}) \neq \Omega \wedge r^{abs}(p_k^{img}, p_l^{img}) \neq \Omega \\ 0 & \text{für } r^{abs}(p_i^{mo}, p_j^{mo}) = \Omega \vee r^{abs}(p_k^{img}, p_l^{img}) = \Omega \end{cases} \quad (45)$$

mit

$$\text{mer}^{abs} = \begin{cases} M^{abs} & \text{für } M^{abs} \geq 0 \\ -1 & \text{für } M^{abs} < 0 \end{cases} \quad M^{abs} = \frac{\max^{\Delta abs} - |r^{abs}(p_i^{mo}, p_j^{mo}) - r^{abs}(p_k^{mo}, p_l^{mo})|}{\max^{\Delta abs}}$$

Darin definieren:

- \max^{disp} Schwellwert für den maximal zugelassenen Abstand zwischen zwei Primitiven.
- $\max^{\Delta azi}$ Schwellwert welcher die maximale Orientierungsdifferenz zwischen zwei Kanten definiert.
- $m^{\Delta len}$ Massstabsfaktor der die maximale Streckung der zurückprojizierten Kante aus dem Modell festlegt.
- $\max^{\Delta abs}$ Schwellwert welcher die maximale Abstandsdifferenz zwischen zwei parallelen Kanten definiert.

Die Funktionen für die Beurteilung der Disparität N^{disp} , der Kantenorientierung N^{azi} und des parallelen Abstandes zwischen zwei Kanten N^{abs} sind auf diese Weise definiert, damit aus einer korrekten Zuordnung ein positiver und aus einer Fehlzuordnung ein negativer Nutzen resultiert. Im Gegensatz dazu resultiert aus der Analyse der Kantenlänge mit der Funktion N^{len} kein oder bei einer Fehlzuordnung ein negativer Nutzen. Entsprechend der vorgeschlagenen Funktion existiert für diesen Fall eine Fehlzuordnung, in welcher die extrahierte Kante aus dem Schlüsselbild bis auf die definierte Toleranz $m^{\Delta len}$ länger ist als die rückprojizierte Modellkante. Die Funktion N^{con} generiert einen positiven Nutzen für Kantenpaare welche in beiden Modellbeschreibungen einen gemeinsamen Stützpunkt aufweisen. Ein negativer Nutzen entsteht lediglich in jenem Fall, bei welchem ein Kantenpaar in \mathbf{m}^{mo} nicht und in \mathbf{m}^{img} verbunden ist. Die vorgeschlagene Definition für die Funktionen N^{len} und N^{con} erklärt sich aus der Tatsache, dass mit dem vorgestellten *high-level* Bildverarbeitungsprozess nicht zwingend das vollständige topologische Modell \mathbf{m}^{img} abgeleitet werden kann. Vielfach ist es nicht möglich, die enthaltenen Kantenprimitiven mit ihrer korrekten Länge aus dem Schlüsselbild zu extrahieren (vgl. dazu Abbildung 3-11). Im Weiteren entsteht definitionsgemäss für eine leere Zuordnung kein Nutzen. Aus den eingeführten Nutzenfunktionen (vgl. Gl. (41) bis (45)) ist nun die eingeführte Evaluationsfunktion für eine Abbildung h wie folgt definiert:

$$\begin{aligned}
 E_h(n) = & \sum_{i=1}^m w_{disp} \cdot N^{disp}(p_i^{mo}, h(p_i^{mo})) + \\
 & \sum_{i=1}^m w_{azi} \cdot N^{azi}(p_i^{mo}, h(p_i^{mo})) + \\
 & \sum_{i=1}^m w_{len} \cdot N^{len}(p_i^{mo}, h(p_i^{mo})) + \\
 & \sum_{i=1}^m \sum_{j=i+1}^m w_{con} \cdot N^{con}(r^{con}(p_i^{mo}, p_j^{mo}), r^{con}(h(p_i^{mo}), p_j^{mo})) + \\
 & \sum_{i=1}^m \sum_{j=i+1}^m w_{abs} \cdot N^{abs}(r^{abs}(p_i^{mo}, p_j^{mo}), r^{abs}(h(p_i^{mo}), p_j^{mo}))
 \end{aligned} \tag{46}$$

Die Faktoren w_{disp} , w_{azi} , w_{len} , w_{con} und w_{abs} definieren Gewichte, mit denen der Einfluss der einzelnen Terme auf den Gesamtnutzen reguliert werden kann.

Optimierung der Evaluationsfunktion

Die optimale Zuordnung zwischen den beiden Modellbeschreibungen m^{mo} und m^{img} ist für dasjenige h gegeben, welches die Evaluationsfunktion maximiert.

$$E_h(n) \Rightarrow \max \tag{47}$$

Für diesen Zweck wird die von Vosselman (1992, Kap. 7) vorgeschlagene informierte Baumsuche eingesetzt. Dieser Optimierungsalgorithmus basiert auf der A* Suche (Pearl, 1984), der für die Optimierung einer Nutzenfunktion modifiziert wurde. Bei diesem Optimierungsansatz wird der Suchraum des Zuordnungsproblems als Baum repräsentiert. Abbildung 3-12 zeigt ein einfaches relationales Zuordnungsproblem (oben) mit dazugehörigem Suchraum (unten). $E(n)$ beschreibt im Kontext der Baumsuche den Gesamtnutzen ausgehend vom Wurzelknoten über den Knoten n bis hin zum Lösungsblattknoten. Die Funktion $E(n)$ kann als Summe der beiden Funktionen $g(n)$ und $h(n)$ angesehen werden, welche entsprechend den anfallenden Nutzen des Pfades vom Wurzelknoten zum Knoten n , bzw. vom Knoten n zum Lösungsblattknoten definieren. Während der Baumsuche kann für den Knoten n der Nutzen $g(n)$ korrekt berechnet werden. Demgegenüber lässt sich zu diesem Zeitpunkt der weitere anfallende Nutzen $h(n)$ lediglich schätzen und wird durch die Funktion $h^*(n)$ bezeichnet. Der geschätzte Gesamtnutzen im Knoten n wird demzufolge wie folgt ausgedrückt:

$$E^*(n) = g(n) + h^*(n) \tag{48}$$

Das Optimierungsproblem wird mit der A* Suche möglichst effizient gelöst, indem immer dieser Knoten im Suchbaum aufgeklappt wird, welcher den höchsten geschätzten Nutzen $E^*(n)$ aufweist. Das Optimum ist erreicht, wenn der aufzuklappende Knoten ein Blattknoten ist und kein Knoten ein höheres $E^*(n)$ besitzt als das berechnete $g(n)$ dieses Blattknotens. Damit das globale Optimum des Zuordnungsproblems korrekt gefunden wird, muss folgende Bedingung

$$h^*(n) \geq h(n) \tag{49}$$

bei der Schätzung des Nutzens erfüllt sein. Je genauer $h(n)$ durch $h^*(n)$ approximiert werden kann, desto schneller kann das Optimierungsproblem gelöst werden. Diese Schätzung von $h^*(n)$ wird auch als eingeführte Heuristik bezeichnet. Wie die Abbildung 3-12 darstellt, kann der Suchraum weiter eingeschränkt werden, indem die Bedingung keine Mehrfachzuordnungen beim Ausklappen eines Knotens berücksichtigt wird. Bei der Berücksichtigung solcher Bedingungen für die Suchraumeinschränkung spricht man auch vom Verfahren des *forward checking*. Der implementierte A* Algorithmus ist im Detail im Anhang F wiedergegeben.

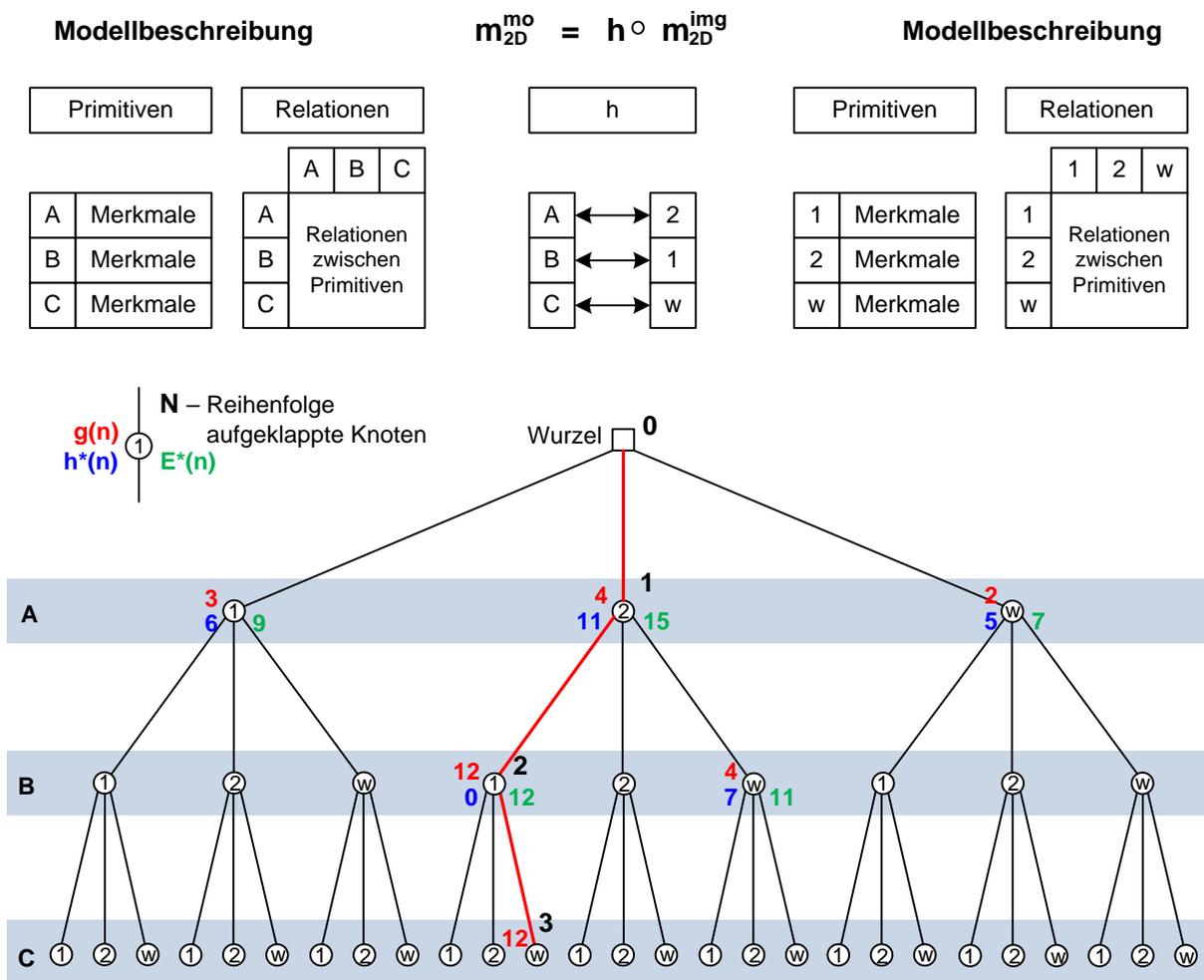


Abbildung 3-12: Informierte Baumsuche zur Lösung einer relationalen Zuordnung

3.3.3 Bildorientierung und Online-Triangulation

Ziel dieses Verarbeitungsschritts ist die unabhängige Bestimmung der äusseren Orientierungsparameter der einzelnen Schlüsselbilder (vgl. Abbildung 3-7). Ausgangslage für die Schätzung ist die vorgängig bestimmte Abbildungsfunktion h zwischen den beiden Modellbeschreibungen \mathbf{m}^{mo} und \mathbf{m}^{img} . Die extrahierten Kanten E_i^{img} aus einem Schlüsselbild werden fortan als Bildbeobachtungen auf die zugeordneten Modellkanten aufgefasst. Die verfügbare 3D-Repräsentation der Modellkanten aus dem digitalen Stadtmodell wird in der Bildorientierung als Referenzgeometrie verwendet. Für die Schätzung der gesuchten Orientierungsparameter aus dem zur Verfügung stehenden Beobachtungsmaterial werden die beiden Ansätze 'Einzelbildorientierung' sowie 'kontinuierlicher Bündelblock' vorgeschlagen. Beide Verfahren basieren auf der perspektivischen Abbildung und werden in der Photogrammetrie klassisch für die indirekte Orientierung eines Einzelbildes bzw. eines Bildverbandes eingesetzt. Beide hier besprochenen Ansätze setzen voraus, dass die innere Orientierung \mathbf{IO} bekannt ist. Im erweiterten Kontext dieser Arbeit können beide Ansätze in Kombination mit der Bild-zu-3D-Stadtmodell-Zuordnung für die automatische absolute Orientierung (Registrierung) beliebig erfasster Einzelbilder und Videosequenzen auf ein zur Verfügung stehendes Referenzmodell eingesetzt werden.

3.3.3.1 Einzelbildorientierung

Die Bildorientierung eines Einzelbildes lässt sich mit Hilfe des räumlichen Rückwärtsschnittes bestimmen. Als Beobachtungen stehen hier im Gegensatz zur klassischen Einzelbildorientierung direkt keine Punktbeobachtungen sondern Geradenbeobachtungen zur Verfügung. Für die Schätzung der äusseren Orientierung wird das von Schwermann (1995) vorgeschlagene Verfahren verwendet, welches zusätzlich die Integration von Geradenbeobachtungen ermöglicht. Zur Stabilisierung der Schätzung werden aus den zur Verfügung stehenden Kantenbeobachtungen, aufgrund der aus der Modellbeschreibung \mathbf{m}^{mo} bekannten Topologieinformation, zusätzliche Punktbeobachtungen für gemeinsame Kantenstützpunkte mittels Geradenschnitt in der Bildebene abgeleitet. Die Schätzung der gesuchten äusseren Orientierungsparameter \mathbf{EO}_{ind} erfolgt mittels Gauss-Newton-Optimierung (vgl. Anhang B.1), welche die resultierenden Verbesserungsquadrate an den Bildbeobachtungen im Falle einer Überbestimmung minimiert. Da das zugrunde gelegte funktionale Modell bezüglich der unbekannt Parameter nicht linear ist, muss dieses an vorliegenden Näherungswerten linearisiert werden (vgl. Gl. (C-2) und (C-3)). Als Näherungswerte werden direkt die zur Verfügung stehende \mathbf{EO}_{app} verwendet (vgl. Abbildung 3-7). Die sechs unbekannt Parameter können geschätzt werden, wenn mindestens drei Punkt- oder Geradenbeobachtungen in allgemeiner Lage aus der automatischen Bild-zu-3D-Stadtmodell-Zuordnung abgeleitet werden können. Zu beachten bleiben jedoch degenerierte Konfigurationen, bei denen die Anordnung der zur Verfügung stehenden Bildbeobachtungen die Orientierungsbestimmung verunmöglichen. Dies tritt

beispielsweise bei Konfigurationen auf, in welchen sämtliche Bildpunktbeobachtungen auf einer Geraden liegen oder sämtliche zur Verfügung stehenden Geradenbeobachtungen zueinander parallel sind.

3.3.3.2 Kontinuierlicher Bündelblock

Dieser Ansatz verfolgt das Ziel die Qualität der zu bestimmenden Orientierungsparameter im Vergleich zur Einzelbildorientierung zu steigern. Weiter sollen die Orientierungsparameter weitestgehend unabhängig vom Erfolg der Bild-zu-3D-Stadtmodell-Zuordnung bestimmt werden können. Für die Erreichung dieser Zielsetzungen wird nachfolgend ein Ansatz vorgeschlagen, der die Verarbeitungskonzepte von SLAM und SfM übernimmt, und diese mit Hilfe photogrammetrischer Algorithmen umsetzt. Dazu wird der von Bayoud (2006) vorgeschlagene Ansatz aufgegriffen und auf die hier vorausgesetzte Monokamerakonfiguration reduziert. SLAM wie auch SfM arbeiten traditionell wie bereits in Abschnitt 2.1.4 eingeführt relativ zu einem lokal eingeführten Referenzsystem. Die vorliegende Problemstellung unterscheidet sich damit vom klassischen SLAM dahingehend, dass hier zusätzlich die fortlaufend kartierte Umgebung und die gleichzeitig geschätzten Orientierungsdaten des Aufnahmesensors auf das bestehende 3D-Stadtmodell registriert werden müssen. Umgesetzt auf das photogrammetrische Vokabular bedeutet dies, dass angelehnt an die zweistufige Orientierung, hier zusätzlich die absolute Orientierung der fortlaufend kartierten Objekte und der Sensororientierungsdaten zu bestimmen sind. Der vorgeschlagene Ansatz schätzt über die letzten n -Schlüsselbilder \mathbf{SB}_{RGB} der Videosequenz die gesuchten Orientierungsparameter und Objektpunkte der Umgebung mit Hilfe einer Bündelblockausgleichung für jedes Schlüsselbild (Update) neu. Die Umsetzung dieses Ansatzes erfordert folgende in Abbildung 3-13 dargestellte Erweiterung in der Architektur der integrierten Georeferenzierung. Als erstes werden automatisch generierte Bildbeobachtungen der zu kartierenden Verknüpfungspunkte (Merkmale der Umgebung) für jedes Schlüsselbild benötigt. Diese werden durch das Modul Merkmalsverfolgung generiert und aktualisieren für jedes Schlüsselbild die Verknüpfungspunktverwaltung. Basierend auf dem von Tomasi und Kanade (1991) vorgeschlagenen KLT-Tracker (Kanade-Lucas-Tomasi *Feature Tracker*) werden automatisch extrahierte Bildmerkmale über jedes Einzelbild der Sequenz verfolgt. Diese für die automatische Verfolgung notwendigen gut geeigneten Bildstellen (Shi und Tomasi, 1994) werden mittels Interest-Operator vorgängig aus dem anliegenden Eingangsbild \mathbf{EB}_{RGB} extrahiert.

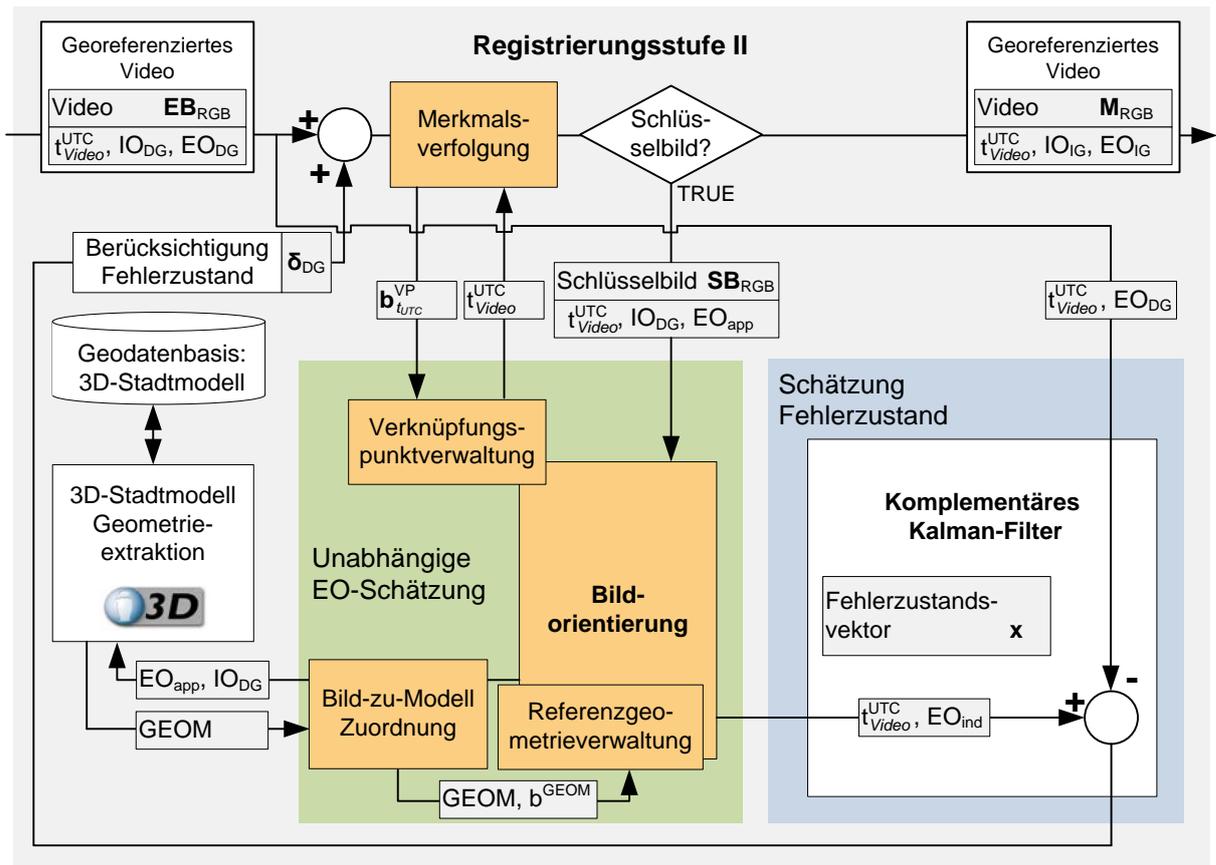


Abbildung 3-13: Architektur Registrierungsstufe II - Integrierte Georeferenzierung für bündelblockbasierte Orientierungsbestimmung

Für die Extraktion wird hier der von Förstner und Gülch (1987) vorgeschlagene Operator verwendet. Durch die zusätzliche Integration einer räumlich unterteilten Merkmalsextraktion im Eingangsbild, wird eine gleichmässige Verteilung der extrahierten Merkmale gewährleistet, welche Voraussetzung ist für eine anschliessende stabil fortlaufende Bündelblocklösung. Die Merkmalsverfolgung extrahiert automatisch neue Bildstellen unter Berücksichtigung der räumlichen Verteilung, wenn eine minimale Anzahl verfolgte Merkmale in der Sequenz unterschritten wird. Beim Anliegen eines Schlüsselbildes SB_{RGB} werden die aktuellen Bildbeobachtungen b^{VP} der verfolgten Merkmale in die Verknüpfungspunktverwaltung transferiert, wo diese der Bündelblockausgleichung zur weiteren Verarbeitung zur Verfügung stehen. Wie eingangs erwähnt ist für die Aufgabenstellung in dieser Arbeit das Ziel, dass die zu kartierenden Merkmale und die Sensororientierungsdaten im Referenzrahmen des vorliegenden 3D-Stadtmodells bestimmt werden. Dieses Ziel wird erreicht, indem die Videosequenz automatisch im geodätischen Bezugsrahmen der vorliegenden Referenzmodellldaten gelagert wird. Zu diesem Zweck wird jeweils das in Abschnitt 3.3.2.2 eingeführte Bild-zu-3D-Stadtmodell-Zuordnungsverfahren für die anliegenden Schlüsselbilder ausgeführt. Resultat dieses Prozesses sind die extrahierten Bildbeobachtungen b^{GEOM} auf zugeordnete Referenzgeometrieelemente **GEOM** für ein Schlüsselbild. Diese Referenzgeometrie-Bildbeobachtungen ermöglichen eine Kartierungs- und Sensororientierungsbestimmung im global gelagerten geodätischen

Referenzkoordinatensystem WGS 84. Für die kontinuierliche Bündelblockschätzung stehen nun sämtliche benötigten Bildbeobachtungen \mathbf{b}^{VP} und \mathbf{b}^{GEOM} sowie für die absolute Lagerung notwendige Referenzgeometrie **GEOM** über die letzten n -Schlüsselbilder zur Verfügung. Damit die unbekanntes Verknüpfungspunktkoordinaten sowie die Sensororientierungsparameter der Schlüsselbilder mit Hilfe einer Bündelblockausgleichung geschätzt werden können, sind Näherungswerte für diese Parameter erforderlich. Der vorgeschlagene Algorithmus sowie das Konzept zur Initialisierung der benötigten Näherungswerte werden im nachfolgenden Abschnitt eingeführt. Aus der kontinuierlichen Bündelblockausgleichung resultieren unabhängig geschätzte Orientierungsparameter \mathbf{EO}_{ind} für das aktuelle Schlüsselbild, aktualisierte Orientierungsparameter der letzten $n-1$ Schlüsselbilder und Verknüpfungspunktkoordinaten der Merkmale. In dieser Arbeit werden die mitgeschätzten Verknüpfungspunkte zurzeit lediglich für die Stabilisierung der visionbasierten unabhängigen Orientierungsbestimmung eingesetzt. Denkbar ist jedoch auch, dass diese als Grundlage für eine automatische detaillierte Umgebungsextraktion aus der Monokamera-Videsequenz im Sinne des SfM-Ansatzes weiter verwendet werden.

Algorithmus und Initialisierung

Abbildung 3-14 zeigt exemplarisch die Aufnahmegeometrie des Bündels a über die letzten n Schlüsselbilder \mathbf{SB}_i ausgehend vom aktuellen Schlüsselbild \mathbf{SB}_a der Sequenz. Ergänzend dazu sind in Abbildung 3-15 die benötigte Datenstruktur und der zugrunde gelegte Algorithmus für den kontinuierlichen Bündelblockausgleich dargestellt. Darin werden folgende Beobachtungstypen für \mathbf{b}^{VP} und \mathbf{b}^{GEOM} verwendet:

\mathbf{b}_i^j	Bildpunktbeobachtung des Schlüsselbildes i auf den Verknüpfungspunkt j .
\mathbf{p}_i^k	Bildpunktbeobachtung des Schlüsselbildes i auf den Punkt k aus dem Referenz-3D-Stadtmodell.
\mathbf{e}_i^l	Bildgeradenbeobachtung des Schlüsselbildes i auf die Kante l aus dem Referenz-3D-Stadtmodell.

Folgende Parameter sind unbekannt und werden in der Ausgleichung geschätzt:

\mathbf{eo}_i^{status}	Äussere Orientierung $(X_0, Y_0, Z_0, \omega, \varphi, \kappa)$ bezüglich des übergeordneten geodätischen Referenzsystems des Schlüsselbildes i .
$^{obs} \mathbf{vp}_j^{status}$	Koordinaten (X_{vp}, Y_{vp}, Z_{vp}) des Verknüpfungspunktes j im übergeordneten geodätischen Referenzsystem. Die Variable obs bezeichnet die Anzahl vorhandener Bildpunktbeobachtungen (Bildstrahlen) auf diesen Verknüpfungspunkt im Bündel a .

Zusätzlich werden folgende Zustände durch die Variable *status* im Algorithmus unterschieden:

- nip* Keine Näherungswerte für die unbekannt Parameter vorhanden (*not initialised parameter*)
- app* Näherungswert für den unbekannt Parameter (*approximation value*)
- adj* Geschätzte Parameter aus der Ausgleichung (*adjusted parameter*)

Die Bündelblockausgleichung erfolgt mittels harter Lagerung direkt auf folgenden Referenzgeometrieelementen **GEOM**:

- rp^k** Koordinaten (X_{RP}, Y_{RP}, Z_{RP}) eines Referenzpunktes *k* aus dem Referenz-3D-Stadtmodell.
- re^l** Geradenbeschreibung ($\alpha_{RE}, \beta_{RE}, \gamma_{RE}, \delta_{RE}$) (Schwermann, 1995) einer Referenzkante *l* aus dem Referenz-3D-Stadtmodell.

Werden nachfolgend die vorgängig eingeführten Bildbeobachtungen, unbekannt Parameter und Referenzgeometrieelemente als Menge aufgefasst, wird dafür zur Beschreibung der jeweilige Grossbuchstabe verwendet. Nach dieser Konvention definiert $\mathbf{B}_i^{1:3} = \{ \mathbf{b}_i^1, \mathbf{b}_i^2, \mathbf{b}_i^3 \}$ die Menge der Verknüpfungspunktbeobachtungen des Schlüsselbildes *i* auf die Verknüpfungspunkte 1,2 und 3.

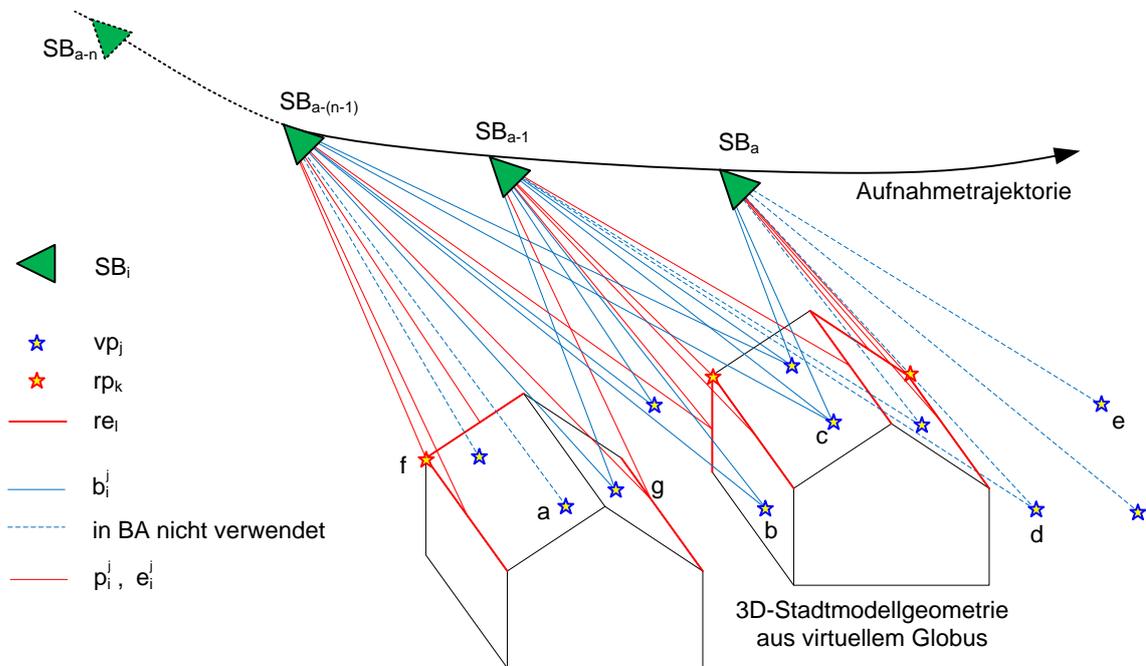


Abbildung 3-14: Aufnahmegeometrie des aktuellen Bündels des Schlüsselbildes *a*

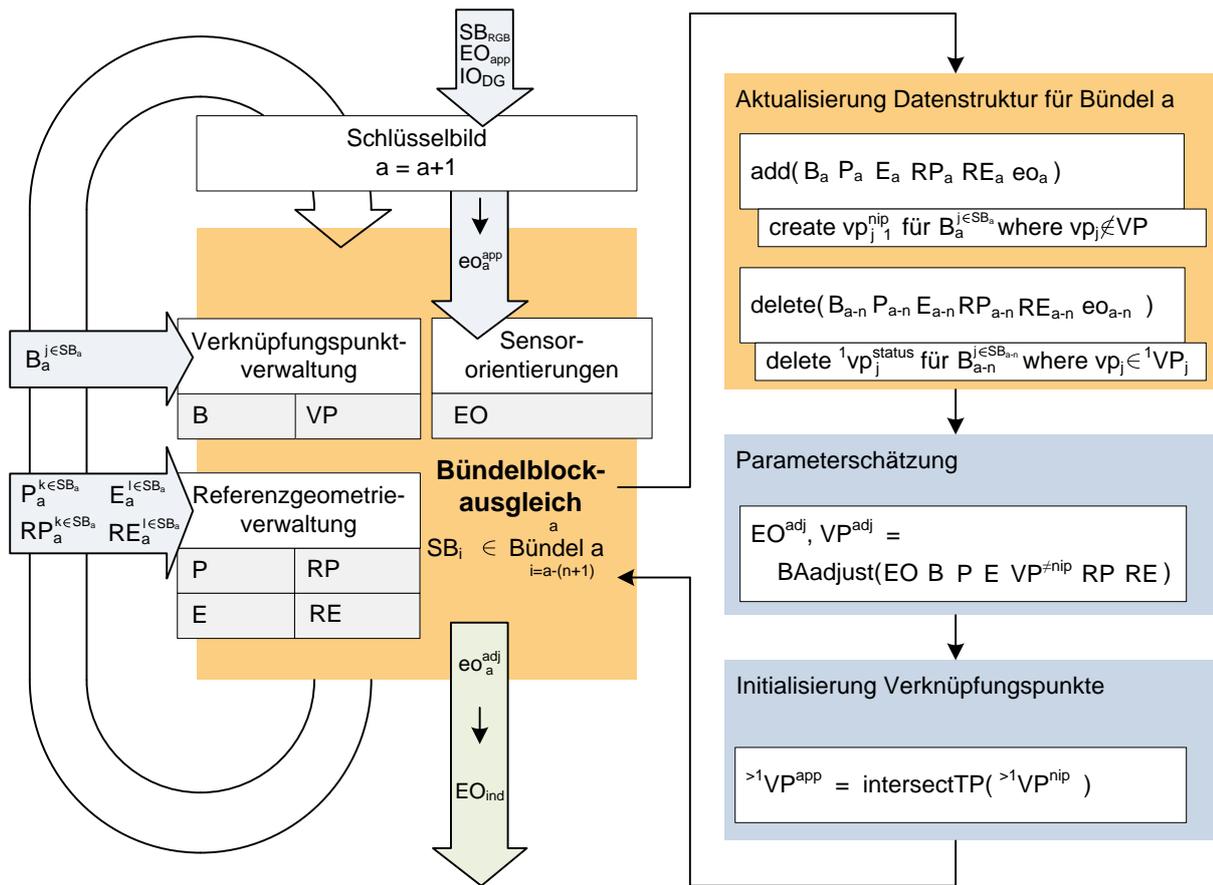


Abbildung 3-15: Architektur und Datenstruktur der kontinuierlichen Bündelblockausgleichung

Liegt ein neues Schlüsselbild für eine unabhängige Orientierungsschätzung vor, werden als erstes die entsprechenden Verknüpfungspunktbeobachtungen $B_a^{j \in SB_a}$ für das aktuelle Schlüsselbild a aus der Merkmalsverfolgung in die Verknüpfungspunktverwaltung integriert. Weiter wird das für das Schlüsselbild zur Verfügung stehende Sensormodell bestehend aus EO_{app} und IO_{DG} für die Bild-zu-3D-Stadtmodell-Zuordnung verwendet. Daraus resultieren für das aktuelle Schlüsselbild a die Referenzgeometriebeobachtungen $P_a^{k \in SB_a}$ und $E_a^{l \in SB_a}$ sowie die zugeordneten 3D-Modell-Geometrielemente $RP_a^{k \in SB_a}$ und $RE_a^{l \in SB_a}$, welche anschliessend in der Referenzgeometrieverwaltung abgelegt werden. Zuletzt wird die äussere Orientierung EO_{app} direkt als Näherungslösung eo_a^{app} in die Bündelblockausgleichung übernommen. Die Bündelblockausgleichung verwendet die Kollinearitätsbeziehung als funktionales Modell, weshalb für die Schätzung der unbekannt Parameter des Bündels a ausreichend genaue Näherungswerte benötigt werden. Als Startwerte in der Schätzung werden zum einen, die aus dem Bündel $a-1$ ausgeglichenen Parameter EO^{adj} und VP^{adj} verwendet. Zum andern stehen für die unbekannt äusseren Orientierungsparameter des Schlüsselbildes a die übernommenen Näherungswerte eo_a^{app} zur Verfügung. Komplexer gestaltet sich jedoch die Berechnung der benötigten Näherungskordinaten VP^{app} von neuen im Bündel auftretenden Verknüpfungspunkten VP^{nip} . Die Initialisierung erfolgt hier nach folgender Strategie.

Die gesuchten Näherungskoordinaten $\geq 2 \mathbf{VP}^{app}$ für $\geq 2 \mathbf{VP}^{nip}$ werden mittels räumlichem Vorwärtsschnitt (Albertz und Wiggenhagen, 2009, Seite 252-253) aus ausgeglichenen \mathbf{EO}^{adj} Schlüsselbildorientierungen bestimmt, für welche mindestens 2 Bildstrahlen vorliegen. Für eine zuverlässige Näherungskoordinatenbestimmung ist eine ausreichend grosse Basis zwischen den verwendeten Schlüsselbildern vorausgesetzt. Liegen für einen Verknüpfungspunkt $> 2 \mathbf{vp}_j^{nip}$ mehr als 3 Bildbeobachtungen vor, erfolgt die Initialisierung aus diesen beiden Schlüsselbildern, welche die grösste zeitliche Differenz aufweisen. Implizit ist hier angenommen, dass aus der kontinuierlichen Kamerabewegung die grösstmögliche Basis resultiert. Basierend auf dem eingeführten Konzept sind die in Abbildung 3-14 mit Buchstaben gekennzeichneten Verknüpfungspunkte und Referenzgeometrielemente wie folgt zu interpretieren:

- a. Initialisierter Verknüpfungspunkt $^1 \mathbf{vp}_j^{adj}$ der im Bündel a nicht mehr mitgeschätzt wird, da nur noch die eine Beobachtung $\mathbf{b}_{a-(n-1)}^j$ vorliegt.
- b. Initialisierter Verknüpfungspunkt $^2 \mathbf{vp}_j^{adj}$ der im Bündel a mitgeschätzt wird, der jedoch nicht bis ins aktuelle Schlüsselbild a verfolgt werden konnte.
- c. Initialisierter Verknüpfungspunkt $^n \mathbf{vp}_j^{app}$ oder $^n \mathbf{vp}_j^{adj}$ welcher aus der maximalen Anzahl an verfügbaren Bildbeobachtungen im Bündel a neu geschätzt wird.
- d. Verknüpfungspunkt $^2 \mathbf{vp}_j^{nip}$ der noch nicht initialisiert ist, jedoch anschliessend an die Bündelausgleichung a mittels Vorwärtsschnitt initialisiert und dem Bündel $a+1$ als $\geq 2 \mathbf{vp}_j^{app}$ zur Verfügung steht.
- e. Neue im Bündel a auftretende Bildbeobachtung \mathbf{b}_a^j auf Verknüpfungspunkt $^1 \mathbf{vp}_j^{nip}$ der weder initialisiert noch anschliessend an den Bündelausgleich a vorwärts eingeschnitten werden kann.
- f. 3D-Modell-Referenzpunkt \mathbf{rp}_a^k mit dazugehöriger Bildpunktbeobachtung $\mathbf{b}_{a-(n-1)}^k$.
- g. 3D-Modell-Referenzkante $\mathbf{RE}_{i \notin a}^l$ mit den beiden Bildgeradenbeobachtungen $\mathbf{E}_{i \notin a}^l$.

Die Initialisierung der neuen \mathbf{VP}^{app} Koordinaten könnte auch vor der eigentlichen Bündelausgleichung a basierend auf der zur Verfügung stehenden Bildorientierung \mathbf{eo}_a^{app} des Schlüsselbildes \mathbf{SB}_a aus der direkten Georeferenzierung erfolgen. Diese Methode hat gegenüber dem vorgeschlagenen Algorithmus den Vorteil, dass ein neuer Verknüpfungspunkt bereits nach zwei statt wie bisher zu drei Schlüsselbildern, in die Bündelblockausgleichung eingeführt werden kann. Untersuchungen zeigen jedoch, dass die erfolgreiche Initialisierung stark von der Genauigkeit der für die Berechnung benötigten Bildorientierungen abhängt. Die direkte Verwendung der näherungsweise bekannten Bildorientierung \mathbf{eo}_a^{app} , bestimmt aus Navigationssensoren geringer Qualität, reicht für diesen Zweck nicht aus, weshalb nur auf Orientierungsdaten aus einer vorgängigen Bündelblockausgleichung bzw. einer Rückwärtsschnittberechnung zurückgegriffen wird. Für interessierte Leser ist der vorgeschlagene Algorithmus der kontinuierlichen Bündelblockausgleichung im Anhang G als Pseudoquelltext dokumentiert.

Parameterschätzung

Die optimale Schätzung der unbekannt Parameter \mathbf{EO} und \mathbf{VP} erfolgt mit dem aus der Photogrammetrie bekannten Verfahren der Bündelblockausgleichung (vgl. Abschnitt

2.3.1.3) auf Basis der zur Verfügung stehenden Bildbeobachtungen \mathbf{B} , \mathbf{P} und \mathbf{E} . Die Datumsfestlegung erfolgt mittels harter Lagerung auf den Referenz-3D-Geometrielementen \mathbf{RP} und \mathbf{RE} . Zusätzlich müssen beim vorgeschlagenen Verfahren folgende zwei Fälle berücksichtigt werden. Das Bündel a kann nicht berechnet werden, wenn ein Datumsdefekt vorliegt. Diese Konstellation ist gegeben, wenn eine ungenügende Anzahl Referenz-Geometrielemente für die Lagerung des Netzes zur Verfügung steht. Im Weiteren kann eine Schlüsselbildorientierung \mathbf{e}_i im Bündel a nicht mitgeschätzt werden, falls dieses durch eine ungenügende Anzahl von Verknüpfungspunkten mit den anderen Bildern verknüpft ist.

Untersuchungen im Zusammenhang mit dem kontinuierlichen Bündelblockausgleich haben gezeigt, dass gerade bei Aufnahmekonfigurationen mit kleiner Basis Konvergenzprobleme mit der klassischen Gauss-Newton-Optimierung auftreten. Aus diesem Grund wird für die Lösung des kontinuierlichen Bündelblockes das gedämpfte nichtlineare Gauss-Newton-Verfahren nach Levenberg-Marquardt (Madsen et al., 1999), im Detail dargestellt in Anhang B.2, verwendet. Zusätzlich wird bei der Auflösung des Normalgleichungssystems die Blockstruktur des Bündelblockproblems ausgenutzt und mit Hilfe des Schurkomplements ein reduziertes Gleichungssystem abgeleitet. Dank diesem lässt sich der Berechnungsaufwand erheblich reduzieren (Engels et al., 2006).

3.3.4 Kontinuierliche Fehlerzustandsschätzung

Wie im Abschnitt 3.3.1 'Architektur des integrierten Lösungsansatzes' eingeführt, schätzt diese Komponente basierend auf der unabhängig bestimmten äusseren Orientierung \mathbf{EO}_{ind} der Schlüsselbilder kontinuierlich den aktuellen Fehlerzustand δ_{DG} der vorliegenden Orientierungsdaten aus der direkten Georeferenzierung \mathbf{EO}_{DG} . Für die Schätzung wird das komplementäre zeitdiskrete Kalman-Filter vorgeschlagen. Nachfolgend werden zuerst einige KF-basierte Integrationsstrategien und Architekturen für die Umsetzung der integrierten Georeferenzierung diskutiert, bevor das bereits in Abschnitt 3.2.2.3 und 3.3.1 eingeführte komplementäre zeitdiskrete Kalman-Filter im Detail eingeführt wird.

3.3.4.1 Filterarchitekturen für die integrierte Georeferenzierung

Für die Umsetzung der vorgeschlagenen integrierten Georeferenzierung sind unterschiedliche Filterarchitekturen und Sensordatenfusionsansätze denkbar. Diese Konzepte und Architekturen können auch unter dem Begriff INS/GNSS/Vision-Sensordatenfusion diskutiert werden. Die aufgezeigte Ausgangslage in Abschnitt 3.2.2.3 sowie die Verfügbarkeit von INS, GNSS und Monokamerasystem legt die Verwendung eines *error state* Kalman-Filters nahe. Mit dieser Architektur werden im Zustandsvektor nicht die gesuchten äusseren Orientierungsparameter selbst, sondern das aktuelle Fehlerverhalten des INS geschätzt, aufgrund dessen die gesuchte verbesserte Navigationslösung berechnet werden kann (Winkler, 2007, Kap. 6) (Shin, 2005). Der *error state* Fusionsansatz lässt sich prinzipiell mit Hilfe eines *tightly coupled* oder *loosely coupled* Kalman-Filters realisieren. Die geschätzten Fehler im Zustandsvektor

können anschliessend mit Hilfe der *feedforward* oder der *feedback* Konfiguration berücksichtigt werden. Die *tightly coupled* Integrationsstrategie arbeitet direkt mit den Rohmessungen der zu fusionierenden Sensoren. Dieser Integrationsansatz erfordert damit den direkten Zugriff auf die Rohbeobachtungen, hier die GNSS-Pseudodistanzen sowie die Bildbeobachtungen in den Schlüsselbildern, für die Schätzung des aktuellen Fehlerzustandes des INS. Im Gegensatz dazu schätzt das *loosely coupled* Filter den gesuchten Fehlerzustand aus bereits im GNSS-Sensor und im Bildorientierungsmodul vorverarbeiteter Positionen und Ausrichtungen. Damit kann bei dieser Integrationsstrategie der Fehlerzustand nur aktualisiert werden, wenn den einzelnen Sensoren während der Vorverarbeitung genügend Beobachtungen für eine vollständige Positions- und Ausrichtungsschätzung zur Verfügung stehen. Die Nutzung bereits abgeleiteter Beobachtungen aus den Rohmesswerten und die nicht immer mögliche Berücksichtigung sämtlicher Rohmesswerte führen im Vergleich zur *tightly coupled* Integration zu einer leicht reduzierten Genauigkeit. Die *feedback* Konfiguration ist gerade in Kombination mit der *tightly coupled* Architektur sinnvoll, da hier die geschätzten Fehlerzustände direkt bei der Lösung der Navigationsrechnung (*strap down* Rechnung) aus den IMU-Rohmesswerten berücksichtigt werden. Im Gegensatz kann die *feedforward* Konfiguration einfach in Kombination mit der *loosely coupled* Integration eingesetzt werden.

Für die Lösung der integrierten Georeferenzierung wird hier folgende angepasste *loosely coupled* Kalman-Filter Architektur mit *feedforward* Fehlerkorrektion vorgeschlagen (vgl. Abbildung 3-7). In der Filterarchitektur wird der aktuelle systematische Fehler δ_{DG} der direkten Georeferenzierungslösung aus den zur Verfügung stehenden unabhängig bestimmten äusseren Orientierungen \mathbf{EO}_{ind} geschätzt. Damit wird im Zustandsvektor \mathbf{x} nicht direkt das Fehlermodell des INS aus GNSS und Monokamera bestimmt, sondern das Fehlerverhalten der vorgängig im Fluglagekontroller aus INS und GNSS fusionierten Positions- und Ausrichtungsdaten aus den zur Verfügung stehenden Schlüsselbildern. Die Verwendung der *error state* Architektur bietet im Vergleich zum *full state* Filter den Vorteil, dass die gesuchten Parameter, hier die äussere Orientierung, unabhängig vom Kalman-Filter zur Verfügung steht und bei erfolgreicher Fehlerzustandsschätzung entsprechend verbessert werden kann. Damit muss im Kalman-Filter nicht das vollständige dynamische Modell des Aufnahmesensors für jedes Einzelbild der Videosequenz geschätzt werden, sondern die Dynamik kann aus der direkten Georeferenzierung übernommen und mit dem aktuellen Fehlerzustand korrigiert werden. Die Aktualisierungsrate des Fehlerzustandes kann dank der gewählten Architektur einfach an die aktuelle Komplexität des Bild-zu-3D-Stadtmodell-Zuordnungsproblems sowie an die Rechenleistung der eingesetzten Hardware anpasst werden. Die vorgeschlagene *loosely coupled* Architektur bietet gerade hinsichtlich der praktischen Umsetzung im Vergleich zu einer IMU-Fehlermodell-basierten *loosely coupled* und *tightly coupled* Integration, bei welchen eine verbesserte Genauigkeit zu erwarten wäre, verschiedene Vorteile. Einerseits kann durch die Fusion bereits vorverarbeiteter äusserer Orientierungen die Navigationssensorik (z.B. Fluglageregelung) einfach ausgetauscht

werden. Andererseits lässt sich diese Sensorfusionsarchitektur einfach in Kombination mit auf verschiedenen Messprinzipien beruhenden Navigationssensoren anwenden, ohne dass die Datenintegration auf Basis der Rohbeobachtungen neu gelöst werden muss. Damit ist die vorgeschlagene Filterarchitektur sehr flexibel einsetzbar und speziell geeignet für die Umsetzung verschiedener Anwendungsszenarien.

3.3.4.2 Komplementäres Kalman-Filter Design

Das Filter-Design basiert auf dem in Brown und Hwang (1997, Kap. 10.1 und 10.4) beschriebenen komplementären zeitdiskreten Kalman-Filter. Der Begriff 'komplementär' beschreibt die Gegebenheit, dass hier die Messungen mehrerer unabhängiger Sensoren fusioniert werden, die dasselbe Signal beobachten. Dabei weisen die Beobachtungen der Sensoren unterschiedliche Rausch-Charakteristiken auf. In dieser Ausprägung des *error state* Filters werden im Zustandsvektor die Fehler der einzelnen Sensoren modelliert. In Abbildung 3-16 ist ein Blockdiagramm des vorgeschlagenen Filters dargestellt.

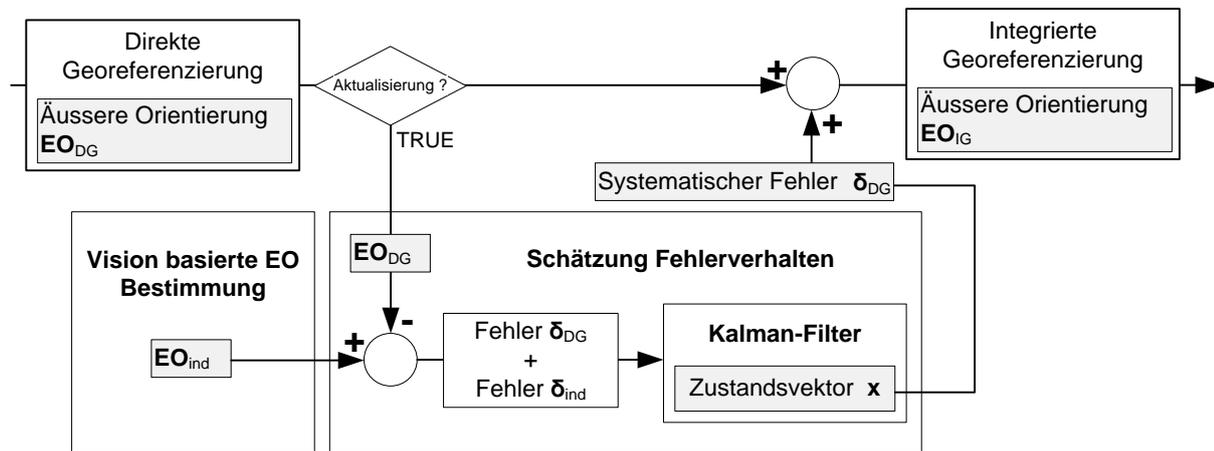


Abbildung 3-16: Komplementäres Kalman-Filter für die integrierte Georeferenzierung

Gesucht ist der aktuelle systematische Fehler δ_{DG} der äusseren Orientierung aus der direkten Georeferenzierung. Der aktuelle systematische Fehleranteil wird aus dem Zustandsvektor \mathbf{x} für jedes Einzelframe prädiziert und am Sensormodell berücksichtigt. Der gesuchte aktuelle Fehler am Projektionszentrum sowie die Fehlauseichnung der direkten Georeferenzierungslösung EO_{DG} werden durch folgenden Vektor

$$\delta_{DG} = (\delta X_0, \delta Y_0, \delta Z_0, \delta \omega, \delta \varphi, \delta \kappa)^T$$

beschrieben. Dieser eingeführte Fehlerterm δ_{DG} fasst die beiden hoch korrelierten Fehleranteile aus der INS/GNSS-Navigationslösung sowie der Systemkalibrierung \mathbf{R}_{gimb}^{img} und $\Delta \mathbf{H}_{body}$ in einem Fehlerterm zusammen. Der eingeführte KF-Systemzustand bildet sich aus dem Zustandsvektor

$$\mathbf{x} = (\delta_{DG}, \dot{\delta}_{DG}, \delta_{ind})^T$$

und der dazugehörigen Schätzfehlerkovarianzmatrix \mathbf{P} und wird aus der Differenz der Stützinformation \mathbf{EO}_{ind} und der äusseren Orientierung \mathbf{EO}_{DG} beobachtet. Darin definiert $\dot{\delta}_{\text{DG}}$ die Änderung der Komponenten des aktuellen systematischen Fehlerterms δ_{DG} . Zusätzlich im Zustandsvektor ist der Term δ_{ind} enthalten, welcher den Fehler der unabhängig bestimmten äusseren Orientierungen \mathbf{EO}_{ind} beschreibt. Aufgrund der in Abschnitt 3.2.2 aufgezeigten Unzulänglichkeiten der Navigationslösung aus Sensoren geringer Qualität wird dem Verhalten des systematischen Fehlers δ_{DG} die integrierte Brownsche Bewegung (engl. *Random Walk*) (Brown und Hwang, 1997, Kap. 2.13) zugrunde gelegt. Demgegenüber ist für δ_{ind} einen Gauss-Markov-Prozess angenommen. Die zusätzliche Mitmodellierung dieses Fehlerterms erlaubt die Berücksichtigung einer allfällig enthaltenen Systematik in der aus den Schlüsselbildern bestimmten Stützinformation. Der Gauss-Markov-Prozess beschränkt im Gegensatz zum *Random Walk* die Grösse des systematischen Fehlers δ_{ind} aufgrund der definierten Standardabweichung σ_{GM} und Prozesslänge β . Die Dynamik dieses Fehlerzustandsmodells wird durch folgendes Differentialgleichungssystem beschrieben:

$$\begin{aligned}\dot{\delta}_{\text{DG}} &= \ddot{\delta}_{\text{DG}} + u_1(t) \\ \ddot{\delta}_{\text{DG}} &= u_2(t) \\ \dot{\delta}_{\text{ind}} &= \beta \cdot \delta_{\text{ind}} + \sqrt{2\sigma_{\text{GM}}^2 \beta}\end{aligned}\quad (50)$$

bei welchem $u_1(t)$ und $u_2(t)$ das weisse Rauschen der Geschwindigkeit bzw. der Beschleunigung definieren. Für das zugrunde gelegte weisse Rauschen $u_1(t)$ und $u_2(t)$ des integrierten *Random-Walk* sowie für die Parameter des verwendeten Gauss-Markov Prozesses σ_{GM} und β werden empirisch bestimmte Werte angenommen. Für die weitere Betrachtung werden die Komponenten im Zustandsvektors \mathbf{x} wie folgt angeordnet

$$\mathbf{x} = (\delta X_{\text{0DG}}, \delta \dot{X}_{\text{0DG}}, \delta X_{\text{0ind}}, \delta Y_{\text{0DG}}, \delta \dot{Y}_{\text{0DG}}, \delta Y_{\text{0ind}}, \delta Z_{\text{0DG}}, \delta \dot{Z}_{\text{0DG}}, \delta Z_{\text{0ind}}, \delta \omega_{\text{DG}}, \delta \dot{\omega}_{\text{DG}}, \delta \omega_{\text{ind}}, \delta \varphi_{\text{DG}}, \delta \dot{\varphi}_{\text{DG}}, \delta \varphi_{\text{ind}}, \delta \kappa_{\text{DG}}, \delta \dot{\kappa}_{\text{DG}}, \delta \kappa_{\text{ind}})^{\text{T}}$$

wodurch das Problem getrennt für jeden einzelnen Parameter der äusseren Orientierung gelöst werden kann. Daraus folgen der Zustandsvektor sowie die dazugehörige Schätzfehlerkovarianzmatrix

$$\mathbf{x}^i = \begin{bmatrix} \delta_{\text{DG}}^i \\ \dot{\delta}_{\text{DG}}^i \\ \delta_{\text{ind}}^i \end{bmatrix}$$

und

$$\mathbf{P}^i = \begin{bmatrix} P_{\delta_{\text{DG}}^i \delta_{\text{DG}}^i} & P_{\delta_{\text{DG}}^i \dot{\delta}_{\text{DG}}^i} & P_{\delta_{\text{DG}}^i \delta_{\text{ind}}^i} \\ P_{\dot{\delta}_{\text{DG}}^i \delta_{\text{DG}}^i} & P_{\dot{\delta}_{\text{DG}}^i \dot{\delta}_{\text{DG}}^i} & P_{\dot{\delta}_{\text{DG}}^i \delta_{\text{ind}}^i} \\ P_{\delta_{\text{ind}}^i \delta_{\text{DG}}^i} & P_{\delta_{\text{ind}}^i \dot{\delta}_{\text{DG}}^i} & P_{\delta_{\text{ind}}^i \delta_{\text{ind}}^i} \end{bmatrix}$$

für jeden Parameter mit $i = X_0, Y_0, Z_0, \omega, \varphi, \kappa$. Damit kann die Lösung des Differentialgleichungssystems (50) in zeitdiskreter Form mit Schrittweite Δt zum Zeitpunkt $t+1$ in abgekürzter Matrixnotation wie folgt geschrieben werden:

$$\underbrace{\begin{bmatrix} \delta_{DG} \\ \dot{\delta}_{DG} \\ \delta_{ind} \end{bmatrix}}_{\mathbf{x}_{t+1}^i} = \underbrace{\begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-\beta\Delta t} \end{bmatrix}}_{\mathbf{\Phi}_{t+1}} \cdot \underbrace{\begin{bmatrix} \delta_{DG} \\ \dot{\delta}_{DG} \\ \delta_{ind} \end{bmatrix}}_{\mathbf{x}_t^i} + \mathbf{w}_t^i \quad (51)$$

Darin definieren $\mathbf{\Phi}_t$ die Transitionsmatrix und \mathbf{w}_t^i das System- oder Prozessrauschen. Aus dem Systemrauschen leitet sich die für die Prädiktion der Schätzfehlerkovarianzmatrix \mathbf{P}_t^i benötigte Autokovarianzmatrix

$$\mathbf{Q}_t^i = \begin{bmatrix} \Delta t \cdot u_1^2 + \frac{\Delta t^3}{3} u_2^2 & \frac{\Delta t^2}{2} u_2^2 & 0 \\ \frac{\Delta t^2}{2} u_2^2 & \Delta t \cdot u_2^2 & 0 \\ 0 & 0 & \sigma_{GM}^2 (1 - e^{-2\beta\Delta t}) \end{bmatrix}$$

aus dem Prozessrauschen ab. Die detaillierte Herleitung dieser Matrizen aus dem dynamischen Modell (50) sind in Brown und Hwang (1997, Kap. 10.1 und 10.4) zu finden.

Zur Vervollständigung des diskreten Kalman-Filters fehlt noch das Beobachtungsmodell, welches für die Aktualisierung des Fehlerzustandes \mathbf{x}_t und \mathbf{P}_t benötigt wird. Aus der in Abbildung 3-16 dargestellten komplementären Filter-Architektur, leiten sich die Beobachtungen \mathbf{z}_t aus der Differenz

$$\mathbf{z}_t = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ \Delta \omega \\ \Delta \varphi \\ \Delta \kappa \end{bmatrix}_t = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ \omega \\ \varphi \\ \kappa \end{bmatrix}_{ind} - \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ \omega \\ \varphi \\ \kappa \end{bmatrix}_{DG} \quad (52)$$

der unabhängig bestimmten \mathbf{EO}_{ind} und der direkt gemessenen äusseren Orientierung \mathbf{EO}_{DG} des Schlüsselbildes zur Epoche t ab. Daraus folgt das Beobachtungsmodell pro Komponente

$$z_t^i = \underbrace{[1 \quad 0 \quad 1]}_{\mathbf{H}^i} \cdot \underbrace{\begin{bmatrix} \delta_{DG} \\ \dot{\delta}_{DG} \\ \delta_{ind} \end{bmatrix}}_{\mathbf{x}_t^i} + v_t^i \quad (53)$$

in welchem die Designmatrix \mathbf{H} den Zusammenhang zwischen Beobachtung und den Komponenten des Fehlerzustandsvektors definiert. Weitergehend beschreibt v_i das Messrauschen der einzelnen Beobachtungen, aus welchem die Kovarianzmatrix

$$\mathbf{R}_t^i = (\sigma_{z_i})^2$$

aufgrund des angenommenen mittleren Fehlers σ_{z_i} jeder Beobachtung resultiert. Der angenommene mittlere Fehler wird direkt aus der Varianz der unabhängig bestimmten äusseren Orientierung abgeleitet. Dabei werden im Falle des kontinuierlichen Bündelblocks, im Gegensatz zur Einzelbildorientierung bei welcher empirisch bestimmte Genauigkeiten verwendet werden, direkt die geschätzten mittleren Fehler übernommen. Damit sind die benötigten Matrizen für den in Abschnitt 2.4 beschriebenen und in Anhang D ausformulierten diskreten Kalman-Filter Algorithmus vollständig.

Das Filter-Design kann dank der Möglichkeit der unterschiedlichen Modellierung der stochastischen Prozesse sehr flexibel eingesetzt werden. Setzt man beispielsweise die Parameter des Gauss-Markov-Prozesses auf Null, resultiert der vereinfachte Fehlerzustand $\mathbf{x} = (\delta_{DG}, \dot{\delta}_{DG})^T$, der direkt mit den unabhängig bestimmten äusseren Orientierungen beobachtet werden kann. Wird andererseits der Parameter $u_1(t) = 0$ gesetzt, resultiert ein integrierter *Random-Walk*, welcher lediglich durch das definierte Rauschen $u_2(t)$ getrieben ist.

3.3.4.3 Berücksichtigung des Fehlerzustandes

Die Aktualisierung des Fehlerzustandes im Kalman-Filter erfolgt für jedes Schlüsselbild. Mit folgender Gleichung kann das aktuelle Fehlermodell der direkten Georeferenzierung auf jedes Einzelbild \mathbf{EB}_{RGB} der Sequenz mit der definierten Fehlerdynamik aus Gleichung (51) für jeden Parameter prädiziert werden.

$$\delta_{DG}^{i, UTC} = \begin{bmatrix} 1 & (t_{\text{Video}}^{UTC} - t) \end{bmatrix} \cdot \begin{bmatrix} \delta_{DG}^i \\ \dot{\delta}_{DG}^i \end{bmatrix}_t \quad (54)$$

Darin definiert t_{Video}^{UTC} Epoche des aktuellen Einzelbildes und t die Epoche der letzten Aktualisierung des Fehlerzustandes. Die gesuchte äussere Orientierung der integrierten Lösung \mathbf{EO}_{IG} für jedes Einzelbild der Sequenz resultiert abschliessend aus der Addition

$$\mathbf{EO}_{IG} = \begin{bmatrix} \mathbf{X}_0 \\ \Theta \end{bmatrix}_{IG}^{t_{\text{Video}}^{UTC}} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ \omega \\ \varphi \\ \kappa \end{bmatrix}_{DG}^{t_{\text{Video}}^{UTC}} + \begin{bmatrix} \delta X_0 \\ \delta Y_0 \\ \delta Z_0 \\ \delta \omega \\ \delta \varphi \\ \delta \kappa \end{bmatrix}_{DG}^{t_{\text{Video}}^{UTC}} \quad (55)$$

$\underbrace{\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ \omega \\ \varphi \\ \kappa \end{bmatrix}_{DG}^{t_{\text{Video}}^{UTC}}}_{\mathbf{EO}_{DG}^{t_{\text{Video}}^{UTC}}}$
 \quad
 $\underbrace{\begin{bmatrix} \delta X_0 \\ \delta Y_0 \\ \delta Z_0 \\ \delta \omega \\ \delta \varphi \\ \delta \kappa \end{bmatrix}_{DG}^{t_{\text{Video}}^{UTC}}}_{\delta_{DG}^{t_{\text{Video}}^{UTC}}}$

des prädizierten Fehlers mit den Komponenten der äusseren Orientierung der direkten Georeferenzierung \mathbf{EO}_{DG} .

4 Prototypenentwicklung und Umsetzung

Dieses Kapitel stellt die Softwarelösung UAVision sowie den im Zusammenhang mit dieser Arbeit realisierten Video-Georeferenzierungs-Prototyp (VGR-Prototyp) vor, welche den Video-Georegistrierungsansatz aus Kapitel 3 implementieren. Anhand der implementierten Programmsysteme wird die Leistungsfähigkeit und Funktionsfähigkeit der eingeführten und vorgeschlagenen Konzepte und Algorithmen im anschließenden Kapitel 'Testflüge und Ergebnisse' untersucht und verifiziert. Der erste Abschnitt führt in die beiden Programmsysteme ein und gibt einen Überblick über die in UAVision und im VGR-Prototyp umgesetzten Konzepte und Komponenten. Darauf aufbauend werden in den folgenden Abschnitten die beiden Programmsysteme, die implementierte Datenverarbeitung sowie die realisierte Funktionalität vorgestellt.

4.1 Umsetzungskonzept und Programmsysteme

Der vorgeschlagene mobile Video-Georegistrierungsansatz wurde in Abschnitt 3.1 im Detail eingeführt und ist in Abbildung 3-1 abgebildet. Die praktische Umsetzung dieser Architektur erfordert die Implementierung von Softwarekomponenten, welche insbesondere die in Abschnitt 3.2 und 3.3 eingeführte direkte- und integrierte Video-Georeferenzierung und die Video-Integration in den virtuellen Globus unterstützen. Parallel zu dieser Arbeit entwickelt die Fachhochschule Nordwestschweiz in Zusammenarbeit mit der SwissCopter AG¹ im Rahmen des Forschungsprojekts UAVision² an der kommerziellen Umsetzung einer drohnenbasierten Video-Georegistrierungslösung. Die in diesem Projekt realisierte Softwarelösung UAVision implementiert die erste Registrierungsstufe und nutzt den virtuellen Globus i3D für die anschließende Echtzeit-Videodatenintegration. Parallel zur UAVision-Softwareentwicklung wurde der VGR-Prototyp realisiert, welcher die vollständige zweistufige Video-Georegistrierung implementiert. Der VGR-Prototyp demonstriert zum einen die Umsetzbarkeit der vorgeschlagenen integrierten Video-Georeferenzierung und diente zum anderen als Testplattform für die Entwicklung der benötigten Algorithmen. Die untenstehende Abbildung 4-1 zeigt zur Verdeutlichung die beiden Programmsysteme und die damit realisierten Komponenten. Im Unterschied zur UAVision-Software können die Videodaten mit dem implementierten VGR-Prototyp nicht in Echtzeit verarbeitet werden. Einerseits eignet sich die verwendete Programmierumgebung nicht für eine Echtzeitdatenverarbeitung und andererseits erlauben die implementierten Algorithmen der zweiten Registrierungsstufe noch keine Echtzeitverarbeitung. Jedoch bietet der VGR-Prototyp die Möglichkeit, das georegistrierte Video mit Hilfe einer zusätzlichen Exportfunktion in UAVision zu integrieren und in dieser zu nutzen.

¹ SwissCopter AG: <http://www.swisscopter.com/de/swisscopter.html>

² KTI-Projekt UAVision: <http://www.fhnw.ch/habg/ivgi/forschung/uavision>

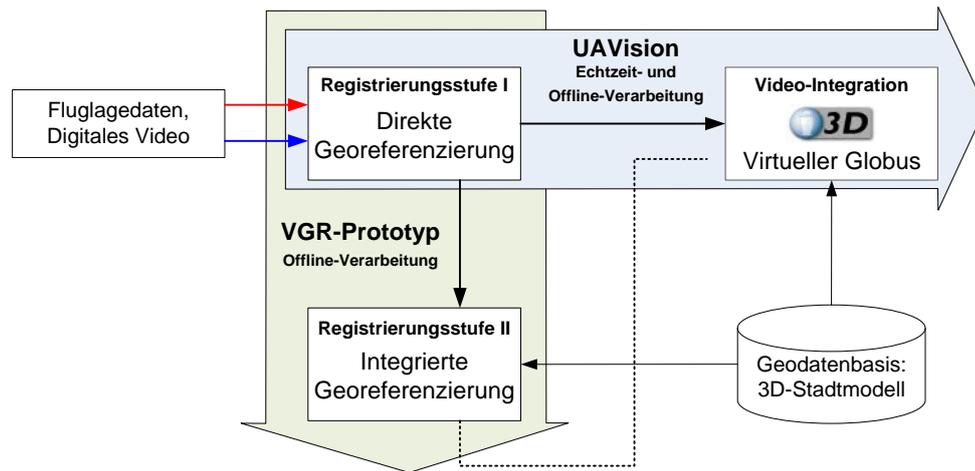


Abbildung 4-1: Umsetzung Video-Georegistrierung und -Integration in UAVision (blau hinterlegt) und im VGR-Prototyp (grün hinterlegt)

4.2 UAVision - Drohnenbasierte direkte Video-Georeferenzierung und -Integration

Die UAVision Software ermöglicht die Planung, Simulation und Durchführung von drohnenbasierten Missionen, sowie die anschließende Echtzeit- oder Offline-Datenintegration der erfassten Bild- und Videodaten in den virtuellen Globus i3D. Die implementierte Funktionalität ist in die Module 'Mission Planning & Simulation', 'Mission Control' und 'Augmented & Virtual Monitoring' unterteilt. Die direkte Video-Georeferenzierungsfunktionalität sowie die anschließende i3D-basierte Video-Integration wurden im Modul 'Augmented & Virtual Monitoring' implementiert, weshalb sich die weiterführenden Betrachtungen im Wesentlichen auf dieses UAVision-Modul konzentrieren. UAVision basiert auf dem virtuellen Globus i3D und erweitert die *Terrain Engine* mit den spezifischen Modulen. Abbildung 4-2 zeigt die UAVision Software Systemarchitektur mit der dazugehörigen graphischen Benutzeroberfläche (GUI - engl. *Graphical User Interface*) des 'Augmented & Virtual Monitoring' Moduls. Weiterführende Informationen zur UAVision-Software sowie zum Forschungsprojekt UAVision im Allgemeinen sind in Eugster et al. (2010) zu finden.

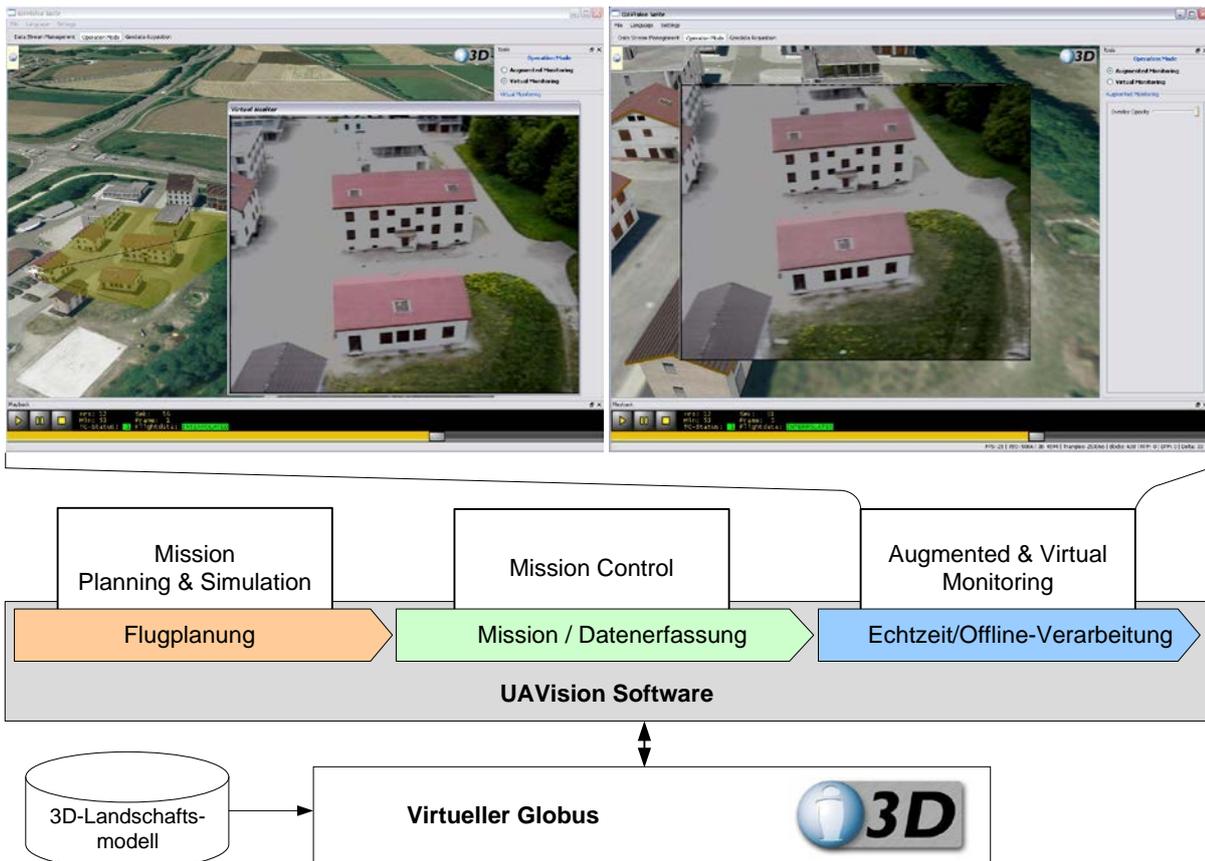


Abbildung 4-2: UAVision Systemarchitektur und GUI (links: 'Virtual Monitoring' Video-Integration, rechts: 'Augmented Monitoring' Video-Integration)

4.2.1 Entwicklungsumgebung

Die UAVision Software wie auch der virtuelle Globus i3D sind ausschliesslich in der Programmiersprache C++ implementiert. C++ eignet sich speziell für die hoch performante Realisierung rechenintensiver Bildverarbeitungsaufgaben und 3D-Computergraphik-Anwendungen. i3D verwendet folgende externen Programmierbibliotheken und -schnittstellen für die Umsetzung folgender Aufgaben:

- Qt Programmbibliothek für die Erstellung von Benutzeroberflächen
- OpenGL Programmiersprachenunabhängige Schnittstelle zur Entwicklung von 2D- und 3D-Computergrafikanwendungen
- OpenCV Programmbibliothek mit Algorithmen für die Video- und Bildverarbeitung
- GDAL Programmbibliothek für die Raster- und Geodatentransformation

UAVision hat vollen Zugriff auf die von i3D zur Verfügung gestellte Funktionalität und die integrierten Programmierbibliotheken. Das Modul 'Augmented & Virtual Monitoring' nutzt OpenCV für den Aufbau der benötigten Video-Player Funktionalität und für die Umsetzung unterschiedlicher Bildverarbeitungsaufgaben.

4.2.2 Video-Datenverarbeitung und -Integration

Das in Abbildung 4-3 gezeigte Sequenzdiagramm erklärt das implementierte Datenverarbeitungskonzept von der direkten Video-Georeferenzierung bis zur Videodatenintegration in eine bestehende 3D-Szene in i3D im Modul 'Augmented & Virtual Monitoring'. Für die Verarbeitung werden drei einzelne *Threads* (bezeichnen einen eigenen Ausführungsstrang in der Abarbeitung eines Programms) verwendet: der Terrain-Engine- oder Haupt-*Thread* (blau), der Video-Player-*Thread* (grün) sowie der Georegistrierungs-*Thread* (orange). Diese werden in der Programmabarbeitung parallel ausgeführt. Der Haupt-*Thread* ist verantwortlich für das Rendering der 3D-Szene in Kombination mit dem aktuell zur Verfügung stehenden Einzelbild aus der Videoquelle. Für die räumlich korrekte Integration des Videobildes wird zusätzlich das dazugehörige Sensormodell **SM** aus der entsprechenden Datenstruktur gelesen und verarbeitet. Unterschiedliche Integrationsansätze sind im folgenden Abschnitt diskutiert. Im Kern der Verarbeitung agiert der Video-Player-*Thread*, welcher Einzelbilder sequentiell aus der angelegten Videoquelle liest und mit Hilfe des *Filter Graphs* verarbeitet. Die Funktionalität des *Filter Graphs* beziehungsweise die benötigten Datenverarbeitungsschritte werden in den einzelnen Filter-Klassen zur Verfügung gestellt. Die realisierten Filter implementieren das in Abschnitt 3.2.1.2 eingeführte direkte Georeferenzierungskonzept. Dabei extrahiert das erste Filter die als Binär-Code integrierte Zeitstempelinformation aus dem aktuellen Einzelbild. Das anschließende Filter bestimmt aufgrund der extrahierten Zeitepoche das gesuchte Sensormodell. Dazu wird das Sensormodell aus den vom Georegistrierungs-*Thread* in der EO-Queue zur Verfügung gestellten äusseren Orientierungen bestimmt. Die Bestimmung erfolgt mittels Interpolation oder Prädiktion sowie der als konstant vorausgesetzten inneren Orientierung. Diese Information wird nun mit dem letzten Filter in die ebenfalls vom Haupt-*Thread* benötigte Datenstruktur gespeichert. Die eigentliche direkte Georeferenzierung erfolgt im Georegistrierungs-*Thread*, welcher aus den anliegenden Fluglagezuständen, den aktuellen Gimballagedaten sowie aus den bekannten Hebelarm- und Fehlaustrichtungskomponenten die vom *Filter Graph* benötigten äusseren Orientierungen bestimmt. Diese Architektur ermöglicht sowohl eine Echtzeit- als auch eine Offline-Verarbeitung. Gerade der Echtzeitfall erfordert das getrennte Lesen und Verarbeiten der unterschiedlich angelegten Datenquellen. Dieses Konzept gewährleistet zum einen, dass der Berechnungsaufwand im Video-Player-*Thread* konstant bleibt, unabhängig von der aktuell vorliegenden Fluglagedaten-Frequenz. Zum anderen wird gewährleistet, dass für die Interpolation oder Prädiktion immer die aktuellsten äusseren Orientierungen zur Verfügung stehen. Im Falle der Offline-Verarbeitung werden mit dem Georegistrierungs-*Thread* fortlaufend Fluglagezustände aus der Fluglagedatei gelesen, sodass für jedes Einzelbild das gesuchte Sensormodell immer aus den äusseren Orientierungen interpoliert werden kann.

rasch zu aktualisieren und damit den Nutzen markant zu steigern. Mögliche Anwendungen reichen unter anderem von der Echtzeit-Überwachung von Infrastrukturanlagen über Waldbrand-Monitoring bis hin zur Entscheidungsunterstützung während und nach Naturkatastrophen.

4.3 VGR-Prototyp - Drohnenbasierte integrierte Video-Georeferenzierung

Der implementierte VGR-Prototyp ermöglicht die Untersuchung der vorgeschlagenen Video-Georegistrierungsansätze und bietet eine ideale Testplattform, um die in Abschnitt 3.2 und 3.3 diskutierten Algorithmen auf ihre Funktionstüchtigkeit hin zu prüfen und zu verbessern. Bei der Implementierung standen in erster Linie die Algorithmenentwicklung im Vordergrund, und noch nicht deren effiziente Realisierung im Hinblick auf eine Echtzeit-Verarbeitung. Dies äussert sich in einem beträchtlichen Potential zur Performancesteigerung der implementierten Algorithmen und Konzepte. Die Prototypanwendung unterstützt die direkte Video-Georeferenzierung sowie die Einzelbildorientierung und die kontinuierliche bündelblockbasierte integrierte Video-Georeferenzierung. Abbildung 4-4 zeigt die umgesetzte Systemarchitektur des VGR-Prototypes.

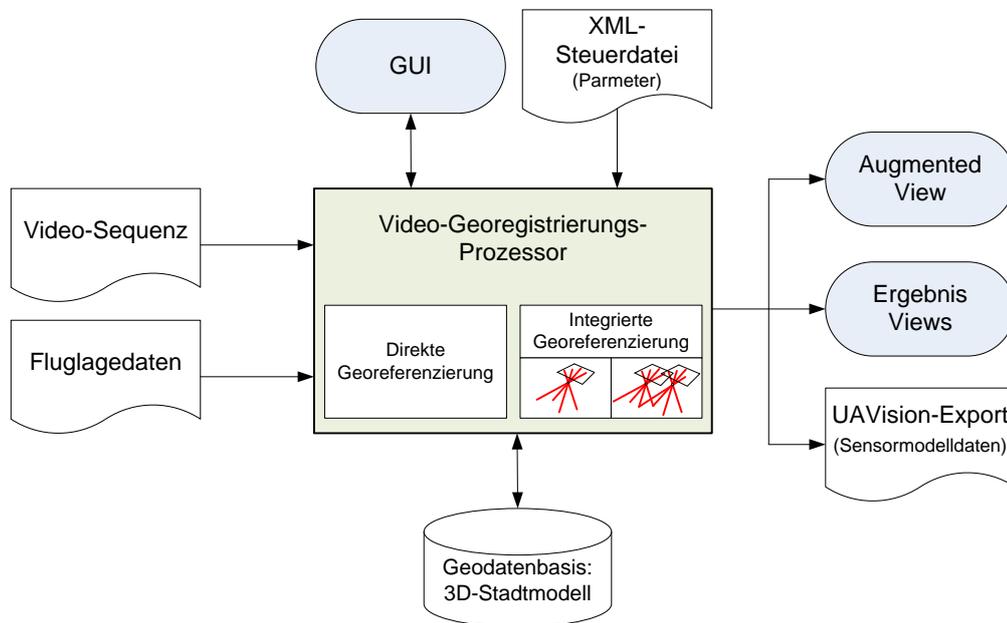


Abbildung 4-4: Systemarchitektur des VGR-Prototypes

Ausgangslage für eine Video-Georegistrierung sind aufgezeichnete Fluglagedaten sowie die dazugehörige Videosequenz. Die Kontrolle der Anwendung erfolgt über eine einfache graphische Benutzeroberfläche. Auf dieser kann eine XML-Steuerdatei (engl. *Extensible Markup Language*) ausgewählt werden, welche die Parameter für die

Datenverarbeitung, die innere Orientierung der verwendeten Videokamera sowie die Hebelarm- und Fehlausrichtungskomponenten festlegt. Diese Architektur erlaubt es auf sehr einfache Weise, für eine Videosequenz unterschiedliche Verarbeitungsparametersätze zu verwalten und zu untersuchen. Die VGR-Prototyp-Implementierung ermöglicht unterschiedliche Datenausgaben. Einerseits kann die aktuelle Lösung aus der direkten und integrierten Georeferenzierung im *Augmented View* gegenübergestellt werden. Andererseits besteht Optional die Möglichkeit diverse weitere Informationen zur Video-Georegistrierung in Form von graphischen Darstellungen und über die Konsole auszugeben. Für eine nachträgliche Verwendung der offline georegistrierten Videosequenz in UAVision steht zusätzlich ein Datenexport zur Verfügung. Die Benutzeroberfläche sowie einzelne mögliche graphische Ergebnisausgaben sind in folgender Abbildung zusammengestellt.

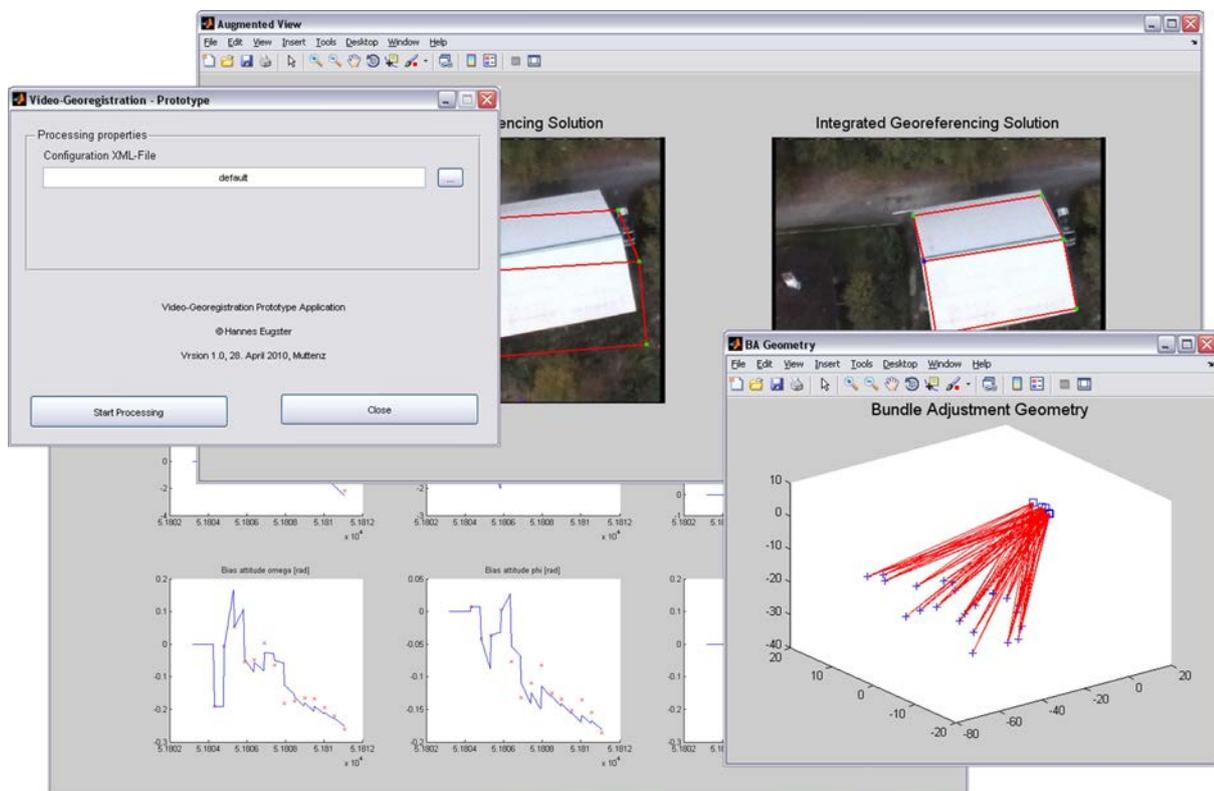


Abbildung 4-5: GUI und graphische Ausgabefenster

4.3.1 Entwicklungsumgebung

Der VGR-Prototyp wurde mit MATLAB¹ (engl. *MATrix LABoratory*) realisiert. MATLAB ist einerseits eine plattformunabhängige Softwarelösung spezialisiert für die numerische Lösung von mathematischen Problemstellungen. Andererseits stellt MATLAB eine proprietäre skriptbasierte Programmiersprache zur Verfügung, mit welcher sehr einfach numerische Berechnungen ausprogrammiert oder Algorithmen entworfen und getestet werden können. Die Programmausführung erfolgt in MATLAB über einen integrierten

¹ Softwareprodukt der Firma The MathWorks, Inc.: <http://www.mathworks.com/>

Interpreter. Für die effiziente Entwicklung stellt MATLAB eine umfangreiche Sammlung von Programmbibliotheken und anwendungsorientierten Werkzeugkisten (engl. *Toolboxes*) zur Verfügung. Besonders interessant und hilfreich für numerische Berechnungen sind die standardmässig zur Verfügung stehenden Algorithmen der linearen Algebra. Der implementierte Prototyp verwendet zusätzlich Funktionen aus den beiden *Toolboxes*:

- Image Processing Programmbibliothek für die Bildverarbeitung
- Mapping Programmbibliothek für die Raster- und Geodaten-
transformation

4.3.2 Umsetzung Video-Georegistrierung

Die Abbildung 4-6 zeigt in Form eines Flussdiagramms die implementierte zweistufige Video-Georegistrierungslösung. Im Unterschied zu UAVision erfolgt hier die gesamte Datenverarbeitung in einem *Thread*. Wie aus dem Diagramm ersichtlich ist, werden die beiden Registrierungsstufen direkte und integrierte Georeferenzierung sequentiell ausgeführt. Die direkte Georeferenzierung ist nach dem in Abschnitt 3.2.1.2 beschriebenen Ansatz implementiert. D.h. die äussere Orientierung wird aus jedem zur Verfügung stehenden Fluglagezustand mit Hilfe der Hebelarm- und Fehlausrichtungskomponenten berechnet und in einer Liste für die Synchronisation mit den Video-Einzelbildern zwischengespeichert. Die direkte Georeferenzierung wird nur ausgeführt, wenn für die Interpolation des gesuchten Sensormodells des aktuellen Einzelbildes der nachfolgende Fluglagezustand benötigt wird. In der anschliessenden integrierten Georeferenzierung werden die beiden Varianten Einzelbildorientierung und kontinuierlicher Bündelblock für die Schätzung des aktuellen Fehlerzustandes der direkt georeferenzierten Lösung unterstützt. Die Aktualisierung des Fehlerzustandes erfolgt für jedes Schlüsselbild der Videosequenz. Zu beachten gilt, dass beim kontinuierlichen Bündelblock Merkmale für die Verknüpfungspunktgenerierung in jedem Einzelbild der Sequenz verfolgt werden (vgl. Abschnitt 3.3.3.2). Als Ausgabe stehen zum einen die *Augmented View* Ausgabe sowie für weitere Datenverarbeitungsschritte die direkt sowie die integriert georeferenzierte Lösung zur Verfügung.

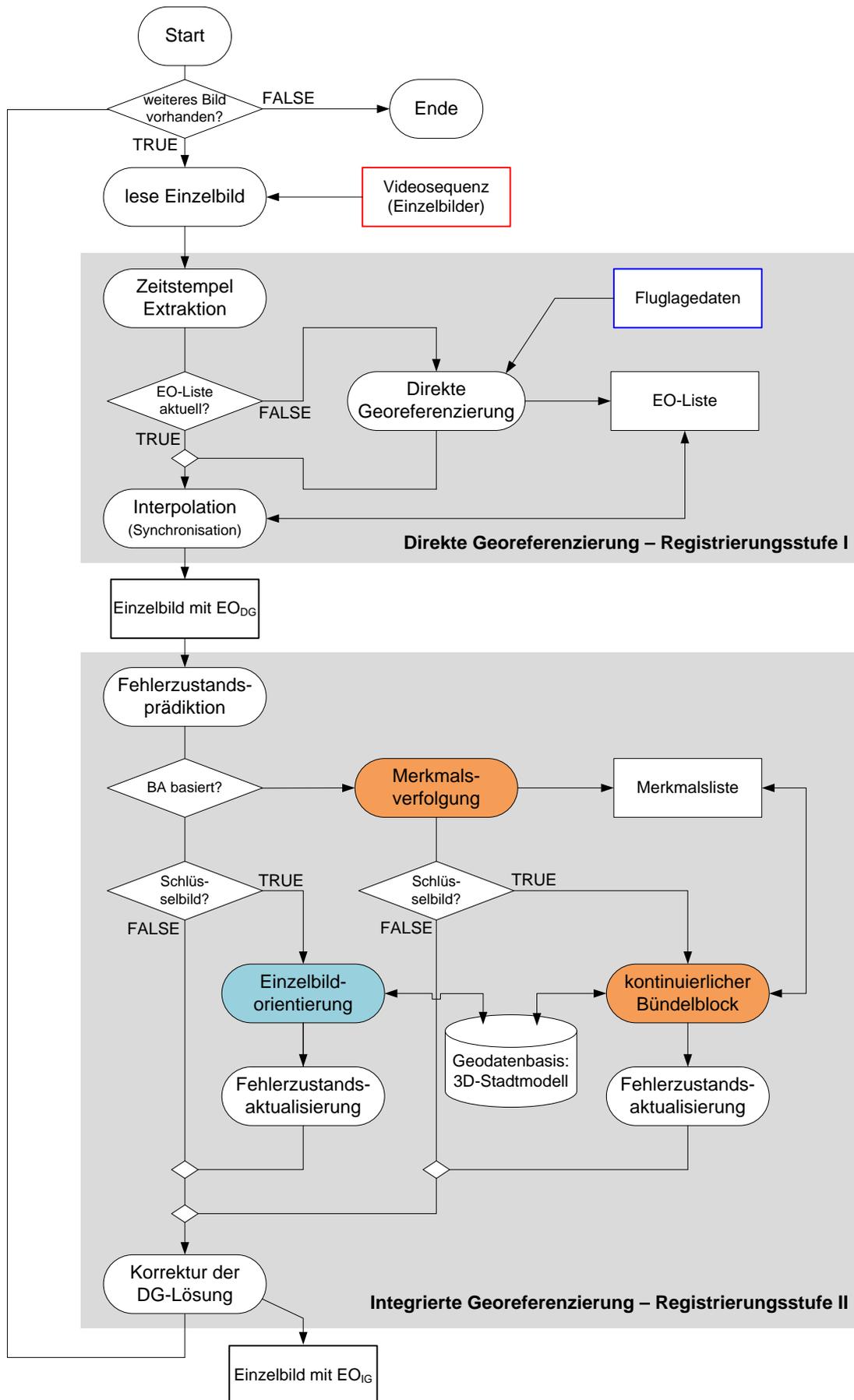


Abbildung 4-6: Flussdiagramm Video-Georegistrierung

4.3.3 Steuerparameter und Prozessierungseinstellungen

Die Definition und Festlegung der aktuellen Steuerparameter und Prozessierungseinstellungen erfolgt über die XML-Steuerdatei. In dieser Steuerdatei können, wie nachfolgende vereinfachte Zusammenstellung aufzeigt, die Eingangsdaten spezifiziert, die Ausgabe festgelegt sowie diverse Prozessierungseinstellungen vorgenommen werden. Gerade im Hinblick auf die Untersuchung und Analyse der implementierten Algorithmen sind sämtliche verwendeten Parameter nach aussen gelegt und können über die Steuerdatei spezifiziert werden.

Eingangsdaten

- Videosequenz Definition der Bildsequenz mit zusätzlicher Angabe des Start- und Endbildes und des Bilddatenformates. Zusätzlich muss die Bildzeile, welche die integrierte Zeitinformation spezifiziert, definiert werden.
- Fluglagedaten Angabe der Fluglagedatendatei.
- 3D-Stadtmodell Referenz auf die für die integrierte Georeferenzierung benötigte Geodatengrundlage.

Ausgabe

- Georeferenzierung Optionale Ausgabe der berechneten äusseren Orientierungsdaten. (Kann anschliessend in UAVision verwendet werden)
- Bildorientierungsprozess Möglichkeit zur Aktivierung zusätzlicher graphischer Ausgaben für die Überwachung der Bild-zu-3D-Modell-Zuordnung sowie des gewählten Bildorientierungsansatzes.

Prozessierungseinstellungen

- Georegistrierung Auswahl der Georegistrierungsstufe: DG oder anschliessend mit IG.
- Innere Orientierung Definition der Videosensorgrosse sowie der Parameter der inneren Orientierung.
- **Direkte Georeferenzierung**
 - Höhenreferenzsystem Angaben zu den aufgezeichneten Höhendaten in den Fluglagedaten. (Geometrischer oder physikalischer Höhenbezug)
 - Gimballagedaten Sind Gimballagedaten vorhanden
 - Fehlausrichtung Definition der Fehlausrichtungskomponenten
 - Hebelarm Definition der Hebelarmkomponenten

- **Integrierte Georeferenzierung**

- Videofrequenz Frequenz der aufgezeichneten Videosequenz.
- Schlüsselbildfrequenz Definition der Schlüsselbildfrequenz für Aktualisierung des Fehlerzustandes.
- Kalman-Filter Kalman-Filter-Einstellungen zur Schätzung des Fehlerzustandes. Parameter zur Definition der stochastischen Prozesse sowie der Genauigkeit der unabhängig bestimmten äusseren Orientierungen.
- Bild-zu-Modell-
 Zuordnung Einstellungen zur relationalen Bild-zu-Modell-Zuordnung. Im Speziellen können die Schwellwerte für die Merkmalsextraktionsalgorithmen (Sobel- oder Canny-Filter und HT) festgelegt werden. Zusätzlich ist festzulegen, ob für die Zuordnung die direkt georeferenzierte oder die integrierte Lösung als Näherung zu verwenden ist.
- Bildorientierungs-
 ansatz Auswahl zwischen 'Einzelbildorientierung' und 'kontinuierlichem Bündelblock' für unabhängige äussere Orientierungsbestimmung.
- Einzelbild-
 orientierung Sollen Punkt- und oder nur Kantenbeobachtungen für die Rückwärtsschnittschätzung verwendet werden.
- Kontinuierlicher
 Bündelblock Detailsinstellungen zur Merkmalsextraktion und Merkmalsverfolgung (Förstner-Operator und KLT-Tracker). Definition der minimalen Anzahl über die Sequenz kontinuierlich verfolgter Merkmale sowie die Anzahl der n-Schlüsselbilder welche in die Bündelblockausgleichung einfließen.

5 Testflüge und Ergebnisse

Anhand unterschiedlicher Videosequenzen wurde das in dieser Arbeit vorgestellte zweistufige Video-Georegistrierungsverfahren hinsichtlich der erreichbaren Georeferenzierungs- und Georegistrierungsgenauigkeit untersucht und beurteilt. Die Erfassung der dafür benötigten Videosequenzen erfolgte mit unbemannten Mikro- und Miniflugplattformen über zwei festgelegten Testgebieten. Für die Beurteilung der diskutierten Ansätze wurden zwei Drohnenkategorien eingesetzt, welche aufgrund ihres Grössen- und Gewichtsunterschieds bezüglich Flugstabilität und äussere Einflüsse wie Wind das erfasste Video unterschiedlich beeinflussen. Weiterführend wurden für eine adäquate Beurteilung der beiden Bildorientierungsansätze 'Einzelbildorientierung' und 'kontinuierlicher Bündelblock' Missionen mit unterschiedlicher Flugdynamik aufgezeichnet. Im ersten Abschnitt werden zuerst die für die Testdatenerfassung zum Einsatz gekommenen UAS sowie die beiden Testgebiete vorgestellt. Anschliessend wird ein Überblick über die erfassten Testsequenzen und Prozessierungskonfigurationen gegeben, gefolgt von einer Zusammenstellung der erreichten Georeferenzierungs- und Georegistrierungsgenauigkeiten und weiteren Untersuchungsergebnissen. Abschliessend werden die vorgeschlagenen Konzepte und Algorithmen sowie der implementierte VGR-Prototyp beurteilt und Erfahrungen diskutiert.

5.1 Aufnahmeplattformen und Testgebiete

5.1.1 Mikro- und Minidrohnensysteme

Die Aufzeichnung der diversen Testvideosequenzen erfolgte mit der Mikrodrohnenplattform *microdrones*¹ md4-200 der Fachhochschule Nordwestschweiz sowie der Minidrohne Copter1B der Firma *weControl AG*². Die beiden Plattformen sind in Abbildung 5-1 abgebildet und die technischen Spezifikationen sind in folgender Tabelle zusammengefasst. Zu beachten ist insbesondere die Positions- und Orientierungsgenauigkeit der INS/GNSS-basierten Fluglagedaten sowie die an der Bodenstation aufgezeichnete Frequenz der übermittelten Fluglagezustände (Telemetriedaten). Die md4-200-Plattform erlaubt zusätzlich die Aufzeichnung von hochfrequenten Fluglagedaten direkt auf einem auf der Plattform integriertem Speicherchip.

¹ *microdrones GmbH*: <http://www.microdrones.com/startseite.php>

² *weControl AG*: <http://www.wecontrol.ch/>



Abbildung 5-1: Mikro- und Minidrohone, microdrones md4-200 (links) - Copter 1B (rechts)

	md4-200	Copter 1B
Plattform		
Hersteller	Microdrones GmbH	SURVEY Copter
Typ	Quadkopter	Rotationsflügler
UAS-Kategorie	Mikro	Mini
Nutzlast	0.2 kg	~5 kg
Max. Abfluggewicht	1 kg	15 kg
Grösse	0.7 m	2 m
Fluglageregelung		
Autopilot	microdrones	wePilot1000
Sensoren für Fluglagezustandsbestimmung	MEMS-INS, GNSS, Baro und Magnetometer	MEMS-INS, GNSS und Barometer
Genauigkeit		
Fluglagedaten		
Lage (CEP)	2.5 m	2.5 m
Höhe (SEP)	5 m	5 m
Nick- und Rollwinkel	1-2° (1 σ)	1° (1 σ)
Azimut	3-5° (1 σ)	1-2° (1 σ)
Fluglagezustandsübertragung	2-4 Hz	-
Fluglagezustandsaufzeichnung	Position: 4 Hz Ausrichtung: 125 Hz	4-5 Hz
PAL-Videokamera		
Typ	Pentax Optio A30 - Video / PAL-Kamera	SONY FCB-EX780BP
Kammerkonstante	~8 mm / ~4.4 mm	~3.2 mm
Auflösung	H: 720, V: 576	H: 720, V: 576

Tabelle 5-1: Technische Spezifikation der eingesetzten Mikro- und Miniplattformssysteme

5.1.2 Testgebiete und 3D-Stadtmodelle

Die diversen Tests und Untersuchungen erfolgten in den beiden Testgebieten 'Ballenberg' und 'ZAZ Eiken'. In beiden Gebieten stehen ausführliche digitale 3D-Stadtmodelle und Geodatengrundlagen zur Verfügung, welche im Rahmen verschiedener Projektarbeiten an der Fachhochschule Nordwestschweiz erfasst worden waren. Im Speziellen das Gebiet ZAZ Eiken eignet sich für vielfältige Tests mit unbemannten Flugsystemen über bebautem Gebiet.

5.1.2.1 Freilichtmuseum Ballenberg

Das Freilichtmuseum Ballenberg¹ befindet sich im Kanton Bern und zeigt traditionelle Gebäude in den unterschiedlichen Baustilen aus sämtlichen Landesteilen der Schweiz. Die einzelnen Gebäude aus den verschiedenen Landesteilen wurden originalgetreu im Freilichtmuseum wieder aufgebaut. Das 3D-Landschaftsmodell des Freilichtmuseums Ballenberg erstellten Geomatikstudierende im Rahmen von Diplom- und Blockkursarbeiten in den Jahren 2006 bis 2009. Die Erfassung und Rekonstruktion sämtlicher Gebäude erfolgte aus einer photogrammetrisch ausgewerteten Dachlandschaft aus hochaufgelösten Luftbildern. Punktuell wurden Gebäude hochdetailliert aus Einzelbildaufnahmen mit Hilfe der Nahbereichsphotogrammetrie und terrestrischem Laserscanning (TLS) rekonstruiert. Zur Vervollständigung des 3D-Landschaftsmodells wurde ein hochaufgelöstes Orthophotomosaik erstellt und das digitale Höhenmodell DTM-AV der swisstopo² übernommen. Für die Erfassung der Einzelbildaufnahmen für die detaillierte Erfassung ausgewählter Gebäude und des Orthophotos kamen die unbemannten Mini- und Mikroplattformen Copter 1B und microdrones md4-200 zum Einsatz. Folgende Abbildung zeigt einen Ausschnitt des 3D-Landschaftsmodells Freilichtmuseum Ballenberg.



Abbildung 5-2: Digitales 3D-Landschaftsmodell Freilichtmuseum Ballenberg

¹ Freilichtmuseum Ballenberg: <http://www.ballenberg.ch/>

² Bundesamt für Landestopografie: <http://www.swisstopo.ch/>

Die rekonstruierten Gebäude stehen in der B-Rep zur Verfügung. Die für die Testflüge verwendeten Gebäudemodelle weisen eine Genauigkeit der Dachlandschaftstützpunkte von ~ 20 cm auf. Auf Basis der neuen Landesvermessung der Schweiz (LV95) lässt sich die 3D-Modellgeometrie verlustfrei in den für die integrierte Video-Georeferenzierung zugrunde gelegten globalen Referenzrahmen WGS 84 transformieren (Swisstopo, 2008).

5.1.2.2 ZAZ Eiken

Das Zivilschutzausbildungszentrum (ZAZ) Eiken im Kanton Aargau stellt Infrastrukturanlagen zur Verfügung für das Einsatztraining von Feuerwehr, Militär, Zivilschutzorganisationen und Rettungsdiensten. Im Wesentlichen besteht dieses Zentrum aus unbewohnten Gebäuden und weiteren für die Nachstellung und Simulation von Katastrophenszenarien benötigten Infrastrukturanlagen. Aus Sicherheitsgründen sind gerade Mini- und Mikrodrohnentestflüge über bebautem Gebiet heikel und nur mit vielen Einschränkungen vertretbar. Das ZAZ bietet aus diesem Grund ideale und annähernd reale Rahmenbedingungen für den Test und die Untersuchung der implementierten Video-Georegistrierungslösung. Im Weiteren ist für den Test der integrierten Georeferenzierung die vorhandene Gebäudesubstanz ideal, da die Gebäude einen durchschnittlichen Dorfkern repräsentieren und sich wegen der einfachen Gebäudegeometrien, Dach- und Fassadenfarben sehr gut eignen für die automatische Kantenextraktion in einer realen Umgebung. Im Rahmen von Forschungsarbeiten im Zusammenhang mit UAS an der Fachhochschule Nordwestschweiz wurde tachymetrisch ein detailliertes 3D-Stadtmodell des Zivilschutzausbildungszentrums erstellt (vgl. Abbildung 5-3). Das digitale Stadtmodell steht ebenfalls in der B-Rep zur Verfügung und weist eine sehr hohe homogene Genauigkeit (~ 10 cm) auf. Die Gebäudegeometrien können ebenfalls verlustfrei in den globalen Referenzrahmen WGS 84 überführt werden.

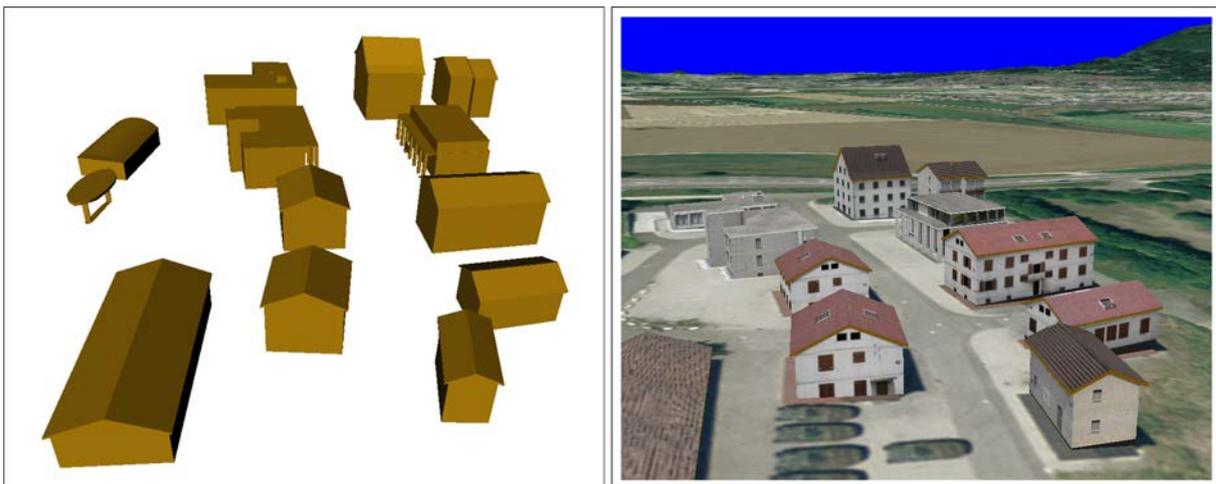


Abbildung 5-3: Digitales 3D-Stadtmodell des Zivilschutzausbildungszentrums Eiken

5.2 Testsequenzen und Untersuchungen

5.2.1 Systemkalibrierung

Voraussetzung für die vorgeschlagene Video-Georegistrierung ist ein kalibriertes Aufnahmesystem wie in Abschnitt 3.1.1 eingeführt. Die Bestimmung der inneren Orientierungsparameter der verwendeten Videokamera sowie der Fehlausrichtung und Hebelarmkomponenten erfolgten für die erfassten Testsequenzen nach den anschliessend beschriebenen Methoden.

5.2.1.1 Innere Orientierung

Für die Schätzung der inneren Orientierungsparameter der Videokamera wurde eine On-the-job-Kalibrierung direkt im Testgebiet durchgeführt. Als räumliches Testfeld diente jeweils ein Gebäude aus dem bestehenden 3D-Stadtmodell mit zusätzlichen am Boden versicherten Passpunkten. Als ideale Objekte erwiesen sich Gebäude, welche einerseits einen grossen Höhenunterschied zwischen Dachgiebel und umliegendem Terrain aufweisen, und andererseits mit der Aufnahmeplattform Bild füllend erfasst werden konnten. Aus ausgewählten Einzelbildern konnte anschliessend die unbekanntere innere Orientierung mit Hilfe einer Bündelblockausgleichung mit zusätzlichen Parametern geschätzt werden. Eine optimale Anordnung der Einzelbilder über dem Testfeld für die zusätzliche Schätzung der gesuchten Parameter der inneren Orientierung (c , x_0 und y_0) ist Luhmann (2003) zu entnehmen. Die Ausgleichung erfolgte mit dem Programmsystem Australis der Firma Photometrix. Für Details zum eingesetzten Kalibrierungsverfahren der inneren Orientierungsparameter ist auf Eugster und Nebiker (2007) verwiesen.

5.2.1.2 Fehlausrichtung und Hebelarm

Damit die erfassten Daten aus GNSS und IMU auf das Projektionszentrum bzw. auf das Kamerakoordinatensystem bezogen werden können, ist die Kenntnis der Anordnung der eingesetzten Sensoren zueinander sowie die Verdrehung zwischen Körperkoordinatensystem und Kamerakoordinatensystem Voraussetzung (vgl. Abschnitt 2.3.2.1). Die Translationen zwischen GNSS-Antennenphasenzentrum, Projektionszentrum und INS-Referenzpunkt wurden direkt terrestrisch eingemessen. Die Fluglageregelung berücksichtigt direkt die Hebelarmkomponente zwischen GNSS-Antennenphasenzentrum und INS. Die Verdrehung zwischen Körperkoordinatensystem und Kamerakoordinatensystem wird zum einen verursacht durch die vom System aufgezeichneten Bewegungen des Gimbals bezüglich dem Körperkoordinatensystem sowie der Fehlausrichtung zwischen Kamerakoordinatensystem und Gimbalssystem. Die Schätzung der Fehlausrichtungswinkel erfolgte aus der aufgezeichneten Videosequenz

eines Kreuzflugs über einem Testfeld durch folgendes Verfahren. In ausgewählten Einzelbildern werden manuell Passpunkte des Testfeldes gemessen. Anschliessend können mit Hilfe einer Gauss-Newton-Optimierung die gesuchten drei Eulerwinkel $\delta\omega, \delta\theta$ und $\delta\psi$ in Gleichung (13) aus den zur Verfügung stehenden Einzelbildern geschätzt werden (Stallmann, 2008).

5.2.2 Testsequenzen und Resultate

Die Untersuchung und Beurteilung der erreichbaren Georeferenzierungs- und Registrierungs Genauigkeit des eingeführten Video-Georegistrierungsansatzes beruht auf folgenden Testsequenzen:

- i. Mittels Minidrohne (Copter 1B) dynamisch (bewegter Aufnahmesichtkegel) erfasste Videosequenz (~200 Einzelbilder, 8 sec) im Testgebiet Freilichtmuseum Ballenberg.
- ii. Mittels Mikrodrohne (md4-200) nahezu statisch (fix bleibender Aufnahmesichtkegel) erfasste Videosequenz (~350 Einzelbilder, 14 sec) im Testgebiet ZAZ Eiken. Für die direkte Georeferenzierung stehen die zur Laufzeit übermittelten und aufgezeichneten Fluglagezustände zur Verfügung.
- iii. Mittels Mikrodrohne (md4-200) nahezu statisch unter starkem Windeinfluss erfasste Videosequenz (~100 Einzelbilder, 4 sec) im Testgebiet ZAZ Eiken. Für die direkte Georeferenzierung stehen die zur Laufzeit übermittelten und aufgezeichneten Fluglagezustände zur Verfügung.
- iv. Gleiche Videosequenz wie in Test **iii**. Im Gegensatz dazu werden hier die auf der Plattform hochfrequent aufgezeichneten Fluglagedaten verwendet.

Bei der Auswertung der erfassten Testsequenzen werden die Prozessierungskonfigurationen

- DG** Direkte Georeferenzierungslösung (vgl. Abschnitt 3.2)
- IG** Integrierte Georeferenzierungslösung (vgl. Abschnitt 3.3)
 - a** Rückwärtsschnittbasierte Verarbeitung (vgl. Abschnitt 3.3.3.1)
 - b** Bündelblockbasierte Verarbeitung (vgl. Abschnitt 3.3.3.2)

unterschieden. Aufgrund der analogen Videodatenübertragung von der unbemannten Flugplattform zur Bodenstation konnten nur sehr kurze ungestörte Sequenzen aufgezeichnet werden. Eine direkte Speicherung des Videos auf der Trägerplattform selbst ist nicht möglich, da aus Gewichtsgründen die Zeitstempelintegration am Boden durchgeführt werden musste (vgl. Abbildung 3-1). Besonders die Prozessierungskonfiguration **IG.b** bedingt störungsfrei aufgezeichnete Videodaten.

5.2.2.1 Genauigkeitsuntersuchung

Für die Untersuchung der erreichbaren Georeferenzierungs- bzw. der Georegistrierungsgenauigkeit werden zum einen unabhängig bestimmte Kontrollpunkte und zum anderen aus der Referenzgeometrie extrahierte Passpunkte verwendet. Der Genauigkeitsnachweis erfolgt aus der Differenz zwischen manuell gemessenen und zurückprojizierten Kontroll- und Passpunkten in beliebigen Einzelbildern der Testsequenz. Die gesuchte Rückprojektion für die ausgewählten Einzelbilder resultiert aus den mit den verschiedenen Prozessierungskonfigurationen berechneten Sensormodellen. Mit den zur Verfügung stehenden Kontrollpunkten wird die erreichte Video-Georeferenzierungsgenauigkeit im übergeordneten Referenzsystem nachgewiesen. Diese Punkte wurden in den Testgebieten signalisiert und mittels relativer GNSS-Positionierung (RTK - engl. *Real-time Kinematic*) mit einer Genauigkeit besser 5 cm eingemessen (Leick, 2004, Kap. 7.9). Für die Untersuchung der erreichten Video-Georegistrierungsgenauigkeit werden direkt natürliche Passpunkte aus der zur Verfügung stehenden Referenzgeometrie verwendet. Die Genauigkeit dieser Punkte liegt im Bereich der angegebenen 3D-Stadtmodellgenauigkeiten der beiden Testgebiete (vgl. Abschnitt 5.1.2). Folgende Tabelle fasst die erreichten Video-Georeferenzierungs- und Georegistrierungsgenauigkeiten der einzelnen Testsequenzen zusammen. Für die Testsequenzen **i**, **iii** und **iv** standen im Untersuchungsgebiet keine Kontrollpunkte zur Verfügung, weshalb lediglich die erreichte Georegistrierungsgenauigkeit ausgewiesen ist.

Testsequenz / Prozessierungs- konfiguration	mittlere Objektdistanz [m]	Georeferenzierungs- genauigkeit		Georegistrierungs- genauigkeit	
		RMS im Bildraum [Pixel]	RMS im Objektraum [m]	RMS im Bildraum [Pixel]	RMS im Objektraum [m]
i / DG	39	-	-	35.0	1.3
i / IG.a	42	-	-	18.4	0.66
i / IG.b	42	-	-	15.9	0.58
ii / DG	53	49.5	3.4	45.5	2.5
ii / IG.a	53	13.5	0.9	13.3	0.7
ii / IG.b	53	27.0	1.9	26.0	1.4
iii / DG	43	-	-	71.0	4.3
iii / IG.a	44	-	-	41.8	2.6
iv / DG	43	-	-	72.3	4.3
iv / IG.a	43	-	-	16.2	1.0

Tabelle 5-2: Erreichte Georeferenzierungs- und Registrierungs-genauigkeiten der untersuchten Testsequenzen

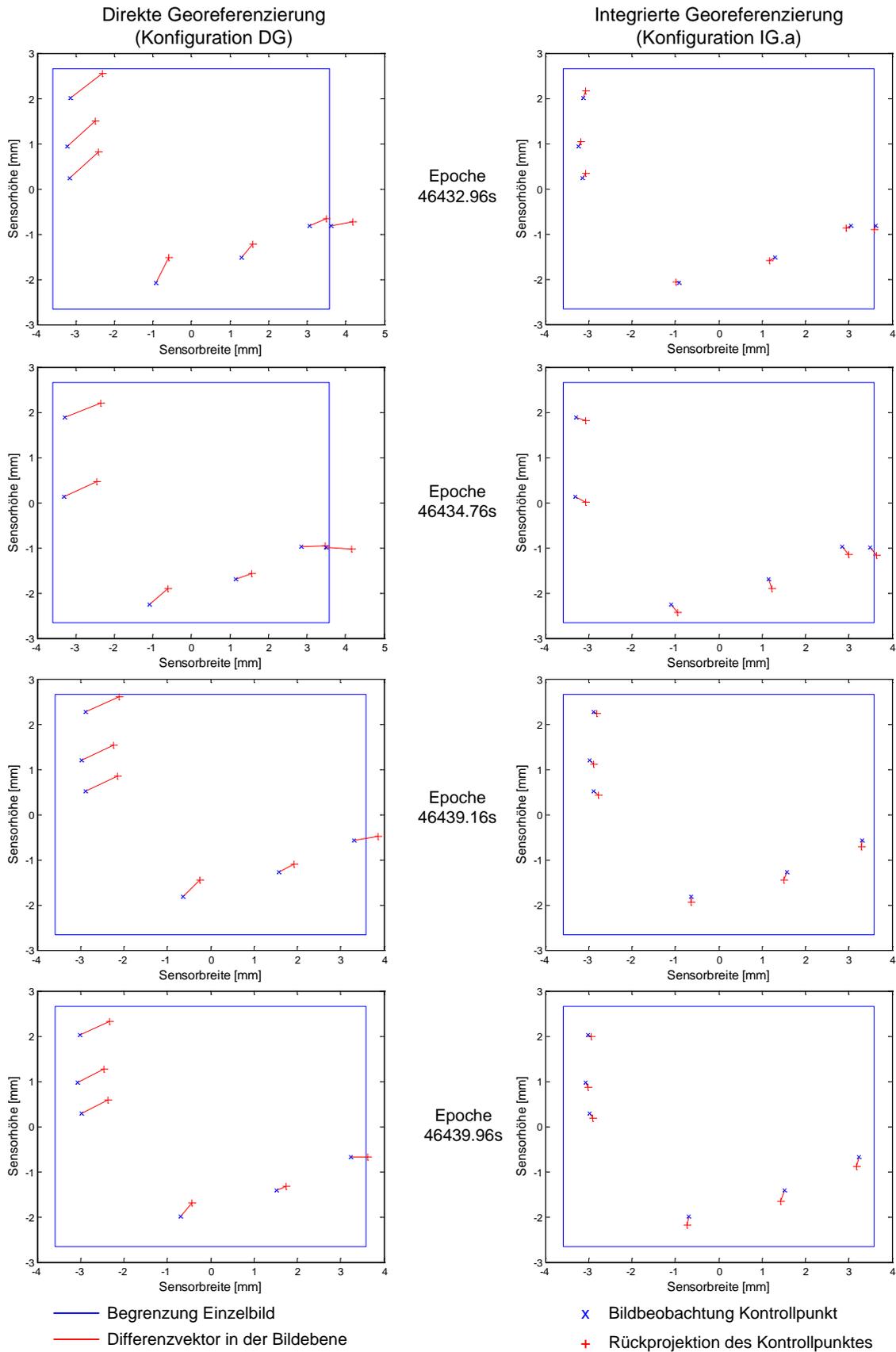


Abbildung 5-4: Fehlerfiguren der Kontrollpunkte der Testsequenz ii, Untersuchte Prozessierungskonfigurationen: Lösung **DG** (links) und **IG.a** (rechts)

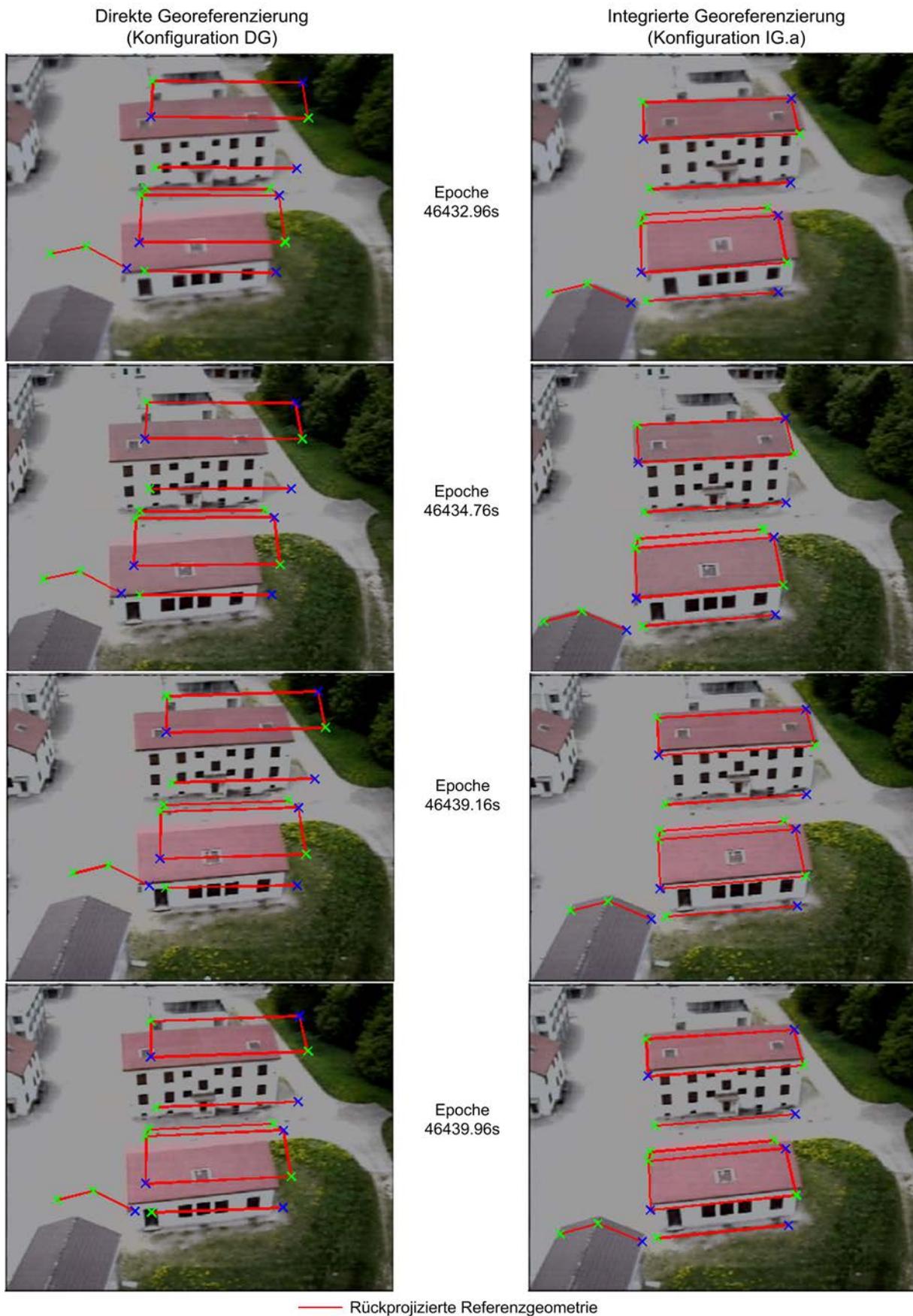


Abbildung 5-5: Georegistrierung Testsequenz ii, Untersuchte Prozessierkonfigurationen: Lösung DG (links) und IG.a (rechts)

5.2.2.2 Genauigkeitsbeurteilung

Die Genauigkeitsuntersuchungen zur direkten Georeferenzierung zeigen eine zeitlich hoch korrelierte verbleibende Systematik. Trotz der vorgängig durchgeführten Fehlausrichtungs- und Hebelarmkalibrierung können die enthaltenen systematischen Effekte nicht kompensiert werden wie dies Abbildung 5-4 und Abbildung 5-5 verdeutlichen (vgl. Diskussion in Abschnitt 3.2.2). Im Weiteren zeigt die mit der Minidrohone aufgezeichnete Sequenz eine signifikant gesteigerte Georegistrierungsgenauigkeit im Vergleich zu den mit der Mikrodrohne erfassten Sequenzen. Erklärt werden kann dieser Effekt in erster Linie mit der stabileren Fluglage der grösseren Aufnahmeplattform sowie der leicht besseren Qualität der zur Verfügung stehenden Fluglagedaten. Die kontinuierliche Schätzung des gesamten Systemfehlerzustandes mit dem vorgeschlagenen integrierten Ansatz ermöglicht eine Genauigkeitssteigerung im Mittel um den Faktor Drei im Vergleich zur direkten Georeferenzierung. Die verbleibenden Fehlerfiguren in Abbildung 5-4 zeigen eine weitgehende Kompensation der in der direkten Georeferenzierung noch enthaltenen systematischen Effekte. Die erreichbare Georegistrierungsgenauigkeit liegt im Submeterbereich unabhängig von der eingesetzten Aufnahmeplattform. Für dynamisch aufgenommene Videosequenzen liefert die bündelblock- im Vergleich zur rückwärtsschnittbasierten integrierten Georeferenzierung leicht bessere Resultate. Diese Aussage darf jedoch nicht auf nahezu statisch erfasste Sequenzen übertragen werden, wie die Resultate der Testsequenz **ii** zeigen. In dieser Situation liefert der bündelblockbasierte Ansatz eine unbefriedigende Lösung, bedingt durch die schwach bestimmte oder nahezu singuläre Netzgeometrie, welche durch die sehr kurzen Basislinien zwischen den einzelnen Schlüsselbildern verursacht wird. Die Resultate der unter starken Windböen aufgezeichneten Testsequenz **iii** und **iv** weisen eine sehr schlechte direkte Georeferenzierungsgenauigkeit auf. Aufgrund der aufgetretenen Windstöße konnte die integrierte Lösung mit den zur Laufzeit übertragenen Fluglagezustände (2-3 Hz) für Testsequenz **iii** nicht vollständig über die gesamte Sequenz initialisiert werden. Probleme treten auf, wenn die benötigten Näherungswerte für die Bild-zu-Modell-Zuordnung aufgrund der Unterabtastung der Plattformbewegung zu ungenau ausfallen. Wie die Resultate der integrierten Lösung der Testsequenz **iv** jedoch dokumentieren, kann dieses Problem mit einer ausreichend hohen Abtastung der Plattformdynamik entschärft werden.

5.2.3 Untersuchungen

5.2.3.1 Kontinuierliche Systemkalibrierung

Die untersuchten Testsequenzen bestätigen die in Abschnitt 3.2.2 aufgezeigten Probleme im Zusammenhang mit Navigationsdaten geringer Qualität. Zum einen zeigt die enthaltene Systematik nach der direkten Georeferenzierung, dass die Positions- und Ausrichtungsfehler der Navigationsdaten die geschätzte Fehlausrichtung und Hebelarmkomponente überlagern. Zum anderen sind die enthaltenen Positions- und Ausrichtungsfehler zeitlich stark korreliert, was sich in den Fehlerfiguren der direkten

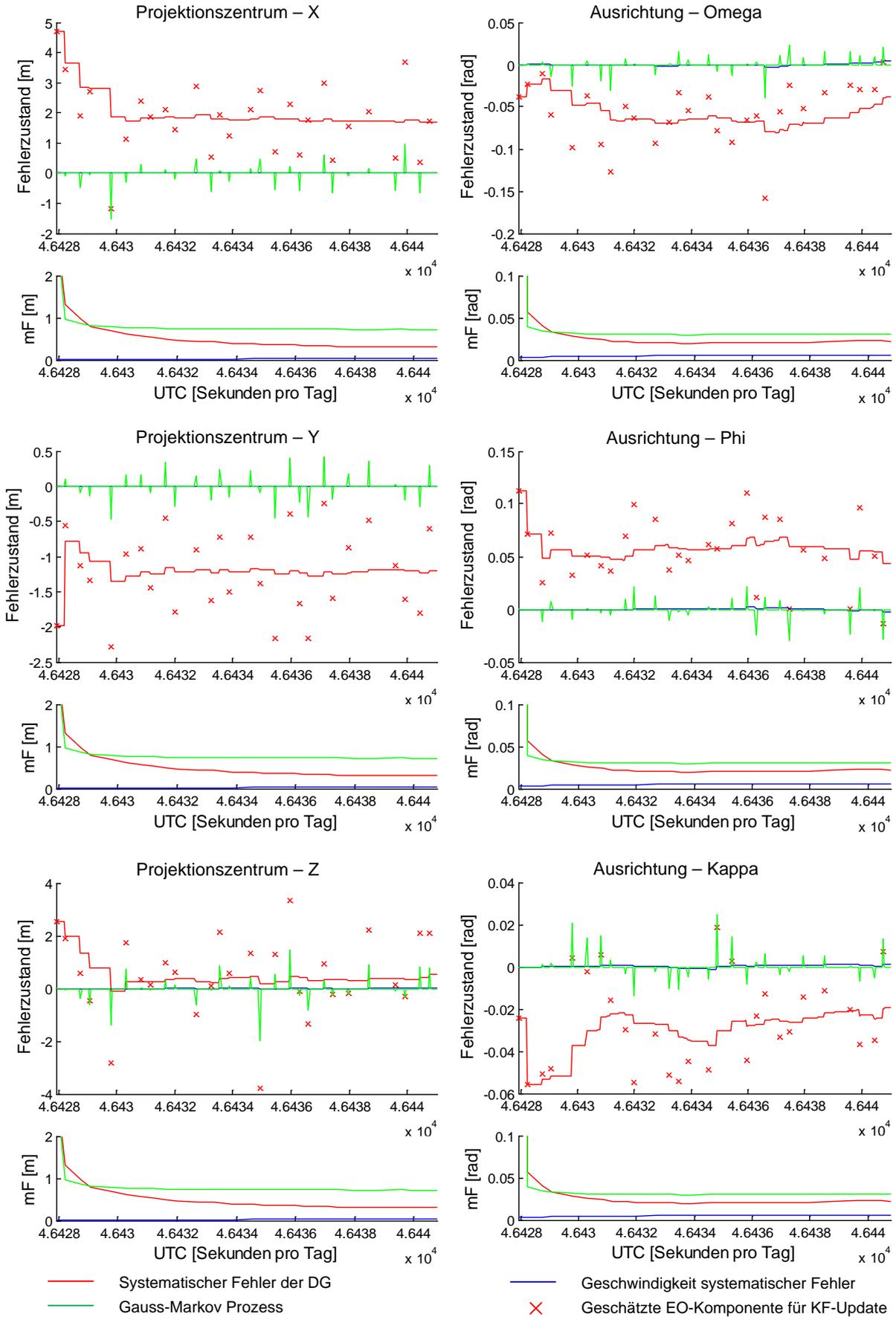


Abbildung 5-6: Geschätzter systematischer Fehler der äusseren Orientierung

Georeferenzierung in Abbildung 5-4 der Testsequenz **ii** manifestiert. Um den sich kontinuierlich ändernden Systemfehler vollständig aufzuzeigen, ist die Testsequenzlänge im Vergleich zur Korrelationslänge der überlagerten Fehleranteile zu kurz. Besser zeigen und bestätigen lässt sich dieser Effekt mit dem diskutierten Verhalten bei der Systemkalibrierung in Abschnitt 3.2.2.1 im Zusammenhang mit Navigationsdaten geringer Qualität. Wie die Genauigkeitsuntersuchungen zur integrierten Georeferenzierung demonstrieren, können mit der vorgeschlagenen kontinuierlichen Fehlerzustandsschätzung die enthaltenen systematischen Effekte weitestgehend geschätzt und kompensiert werden. Abbildung 5-6 zeigt die während der integrierten Verarbeitung im komplementären Kalman-Filter geschätzten Parameter des Zustandsvektors \mathbf{x} – getrennt dargestellt für jede Komponente der äusseren Orientierung – der Testsequenz **ii** / **IG.a**. Darin repräsentiert die rote Kurve den geschätzten und anschliessend kompensierten systematischen Fehleranteil der direkten Georeferenzierungslösung. Die roten Kreuze zeigen die unabhängig bestimmten äusseren Orientierungsparameter für jedes Schlüsselbild welche in den Kalman-Filter *Update* einfließen. Im unteren Abbildungsteil (mF) ist jeweils der mittlere Fehler der geschätzten Parameter des Zustandsvektors dargestellt. Auf weitere in der integrierten Lösung enthaltene und nicht kompensierte systematische Effekte wird in Abschnitt 5.2.3.2 detailliert eingegangen.

5.2.3.2 Fluglagezustandsfrequenz und -Qualität

Wie die Untersuchungen in Abschnitt 5.2.2.1 zeigen, hängt die erreichbare Genauigkeit ab von der verfügbaren Frequenz der Fluglagedaten, der Qualität der Navigationsensoren sowie von der eingesetzten Aufnahmeträgerplattform. Die Fehlerfiguren der integrierten Lösung in Abbildung 5-4 zeigen eine bestehende Restsystematik. Auf der einen Seite hängen diese Fehler von der stochastischen Modellierung des Prozessrauschens und der angenommenen Kovarianzmatrix der unabhängig bestimmten Orientierungsupdates im komplementären Kalman-Filter ab. Auf der anderen Seite kann dieser Effekt, besonders der hochfrequente Anteil, mit der nicht vollständig abgebildeten Aufnahmeplattfordynamik in den zur Verfügung stehenden Fluglagedaten erklärt werden (vgl. dazu Diskussion in Abschnitt 3.2.1.2). Dieser Effekt lässt sich anhand der Genauigkeitsuntersuchung der Testsequenz **iv** im Vergleich zu Sequenz **iii** (vgl. Tabelle 5-2) nachweisen. Die Genauigkeit der integrierten wie auch der direkten Georeferenzierungslösung lässt sich mit Hilfe der auf der Plattform in hoher Frequenz aufgezeichneten Orientierungsdaten verbessern. Im Vergleich zur Testsequenz **iii** werden damit hochfrequente Bewegungen der Aufnahmeplattform erfasst und in der Video-Georegistrierungsprozessierung berücksichtigt. Zusätzlich wird damit die kontinuierliche Systemfehlerzustandsschätzung durch Berücksichtigung dieses Effekts verbessert. Gerade Mikroplattformsysteme sind im Vergleich zu Minidrohnen sehr Windanfällig, weshalb dieser Effekt bei dieser Plattformkategorie stärker ausgeprägt ist. Wie die Untersuchungen zur Testsequenz **i** aufzeigen, reicht eine Fluglagezustandsübertragung von 4-5 Hz für Minidrohnen im Vergleich zu Mikroplattformsystem für eine optimale Georeferenzierungs- und Registrierungsgenauigkeit aus. Weiter wird die

enthaltene Restsystematik verstärkt durch das in der Navigationslösung von INS/GNSS-Sensoren geringer Qualität enthaltene – und durch das komplementäre Kalman-Filter nicht kompensierte – Rauschen (vgl. Abbildung 3-5).

5.2.3.3 Einfluss Referenzgeometriegenauigkeit

Eine offene Frage bleibt, wie wirkt sich die Referenzgeometriegenauigkeit auf die erreichbare Georeferenzierungs- bzw. Georegistrierungsgenauigkeit aus. Für die Untersuchung dieser Fragestellung wird die vorliegende 3D-Stadtmodellreferenzgeometrie wie folgt künstlich verändert:

- a Modellierung eines definierten Rauschens σ in den Stützpunkten der B-Rep
- b Anbringen einer definierten systematischen Verschiebung \mathbf{dB} an der Referenzgeometrie

Situation **a** entspricht der Annahme, dass die vorliegende 3D-Stadtmodellreferenzgeometrie mit einer entsprechenden geringeren Messgenauigkeit erfasst und modelliert ist. In Situation **b** wird angenommen, dass die 3D-Stadtmodellgeometrie konsistent in sich vorliegt, jedoch bezüglich des übergeordneten geodätischen Referenzsystems systematisch verschoben oder verdreht ist. Diese Situation tritt beispielsweise auf, wenn der Datumsübergang zwischen globalem geodätischem Referenzsystem (z.B. WGS 84) und lokalem geodätischem Referenzsystem nicht streng definiert ist. In Praxisdatensätzen können beide Situationen in Kombination auftreten. Für die Untersuchungen wurde Testsequenz **ii** und die Prozessierungskonfiguration **IG.b** verwendet.

Situation / Fehlermodell	mittlere Objektdistanz [m]	Georeferenzierungsgenauigkeit		Georegistrierungsgenauigkeit	
		RMS im Bildraum [Pixel]	RMS im Objektraum [m]	RMS im Bildraum [Pixel]	RMS im Objektraum [m]
a / $\sigma = 0.1$ m	59	32.3	2.5	27.5	1.7
a / $\sigma = 0.2$ m	58	26.3	2.0	21.8	1.3
b / \mathbf{dB} = [0.5 1 0.5] m	53	15.0	1.0	13.5	0.7
b / \mathbf{dB} = [1 4 2] m	55	22.2	1.6	12.2	0.7

Tabelle 5-3: Einfluss der Referenzgeometriegenauigkeit

Die Resultate zeigen einen signifikanten Unterschied zwischen der erreichbaren Georeferenzierungs- und der Georegistrierungsgenauigkeit besonders für die spezifizierte Situation **b** und begründet damit die Unterscheidung dieser beiden Genauigkeitsbegriffe. Die integrierte Georeferenzierung erlaubt die Realisierung einer

korrekten Video-Georegistrierung auf eine bestehende Referenzgeometrie, unabhängig von der Genauigkeit der vorliegenden 3D-Stadtmodelldaten bezüglich eines geodätischen Referenzsystems. Diese Tatsache ist gerade im Zusammenhang mit der Umsetzung einer Augmented Reality Anwendung von grossem Interesse. Wie die Untersuchung zeigt, ist hingegen die Georeferenzierungsgenauigkeit in **b** mit der enthaltenen Systematik zwischen 3D-Stadtmodell und geodätischem Referenzsystem überlagert. Die erreichten Georeferenzierungs- bzw. Georegistrierungsgenauigkeiten für die Situation **a** zeigen, dass zur Ausschöpfung des möglichen Genauigkeitspotentials der integrierten Lösung die relative Referenzgeometriegenauigkeit 20 cm nicht überschreiten sollte. Im Speziellen Nichtparallelitäten durch verrauschte Stützpunkte in der B-Rep haben entscheidenden Einfluss auf die unabhängig bestimmten äusseren Orientierungsparameter.

5.2.3.4 Unabhängige Orientierungsupdatebestimmung

Wie Abschnitt 5.2.2.1 aufzeigt, resultieren für dynamisch oder nahezu statisch aufgezeichnete Videosequenzen mit dem kontinuierlichen Bündelblockansatz und dem Einzelbildorientierungsansatz unterschiedliche Genauigkeiten. Prinzipiell darf von der kontinuierlichen bündelblockbasierten Verarbeitung, aufgrund der verbesserten Aufnahmegeometrie durch die zusätzliche Verknüpfung mehrerer Schlüsselbilder, eine gesteigerte Genauigkeit erwartet werden. Dies zeigen die Resultate der dynamisch aufgenommenen Testsequenz **i** in Abbildung 5-7 und Abbildung 5-8. Die rückwärts-schnittbasierte unabhängige Orientierungsschätzung zeigt ein höheres Rauschen der Projektionszentren und Eulerwinkel im Vergleich zum bündelblockbasierten Ansatz. Damit zeigt sich, dass der kontinuierliche Bündelblockansatz weit weniger beeinflusst ist, von den im aktuellen Schlüsselbild vorliegenden Referenzgeometriebeobachtungen. Im Speziellen eine schlechte Verteilung der Referenzgeometriebeobachtungen im Bild hat vielfach eine instabile Rückwärtsschnittlösung zur Folge (Schwermann, 1995). Der bündelblockbasierte Ansatz ist zudem in der Lage, auch bei Versagen des Bild-zu-Modell-Zuordnungsprozesses, aufgrund der initialisierten Verknüpfungspunkte eine unabhängige äussere Orientierungsbestimmung zu gewährleisten, wie die Testsequenz **i** beispielsweise am Ende aufzeigt. Einzige Bedingung für die erfolgreiche Berechnung einer unabhängigen Bildorientierung ist diesbezüglich eine gültige Datumsfestlegung im Bündel. Damit ist die Beobachtbarkeit des Fehlerterms im zeitdiskreten Kalman-Filter weit weniger abhängig vom sehr rechenintensiven und komplexen Bild-zu-3D-Stadtmodell-Zuordnungsverfahren. Jedoch hat der vorgestellte kontinuierliche Bündelblock-Algorithmus für nahezu statisch erfasste Videosequenzen (kleine Basis zwischen Schlüsselbildern) Nachteile. In diesem speziellen Fall gestaltet sich die Initialisierung der Verknüpfungspunkte sehr schwierig, was in der derzeitigen Implementierung eine unabhängige Orientierungsbestimmung für das aktuelle Schlüsselbild verunmöglicht. Das resultierende Normalgleichungssystem ist aufgrund der statischen Aufnahmegeometrie schlecht konditioniert, was zu Konvergenzproblemen oder zur Singularität führen kann.

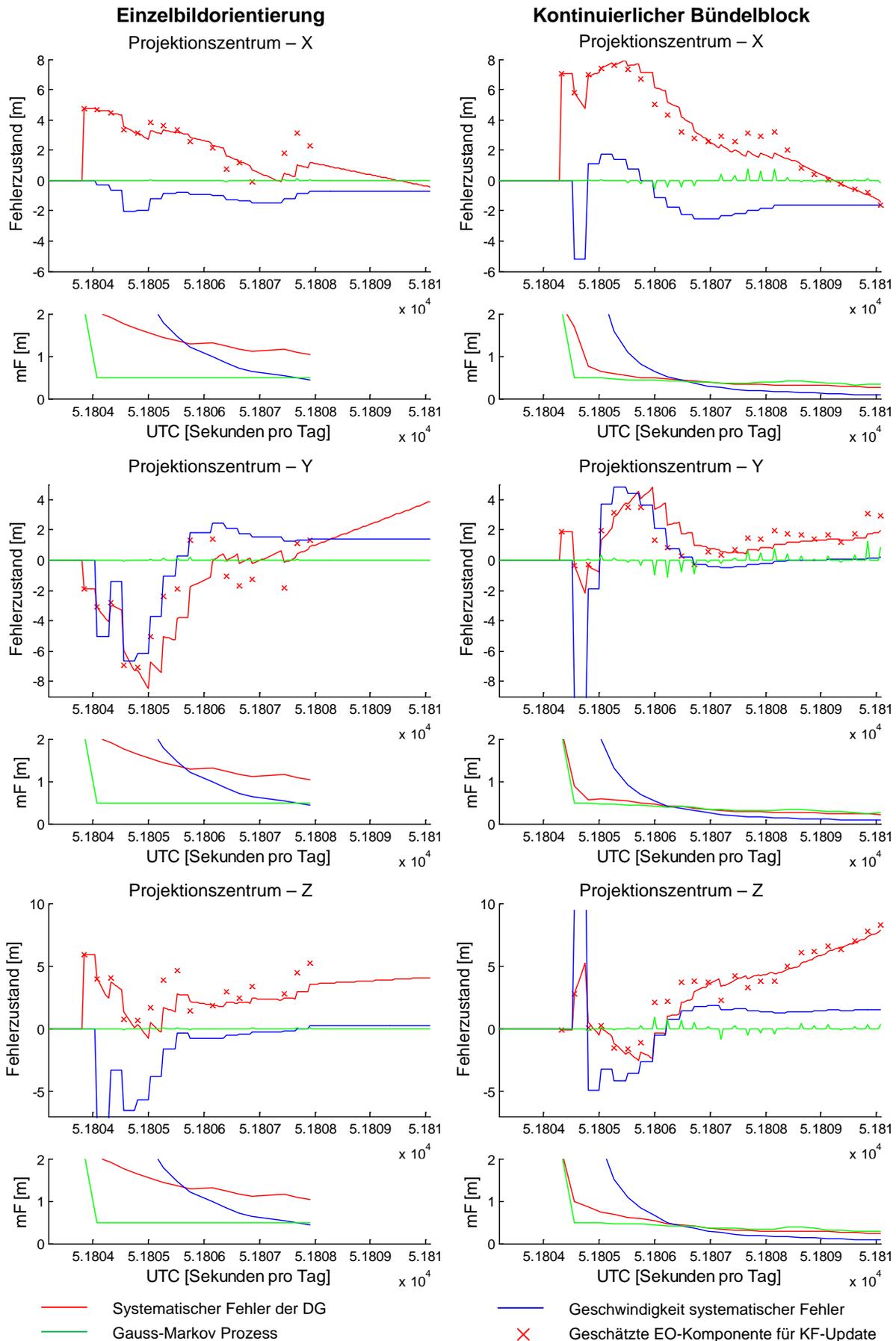


Abbildung 5-7: Äussere Orientierung Testsequenz Ballenberg (Projektionszentrum)

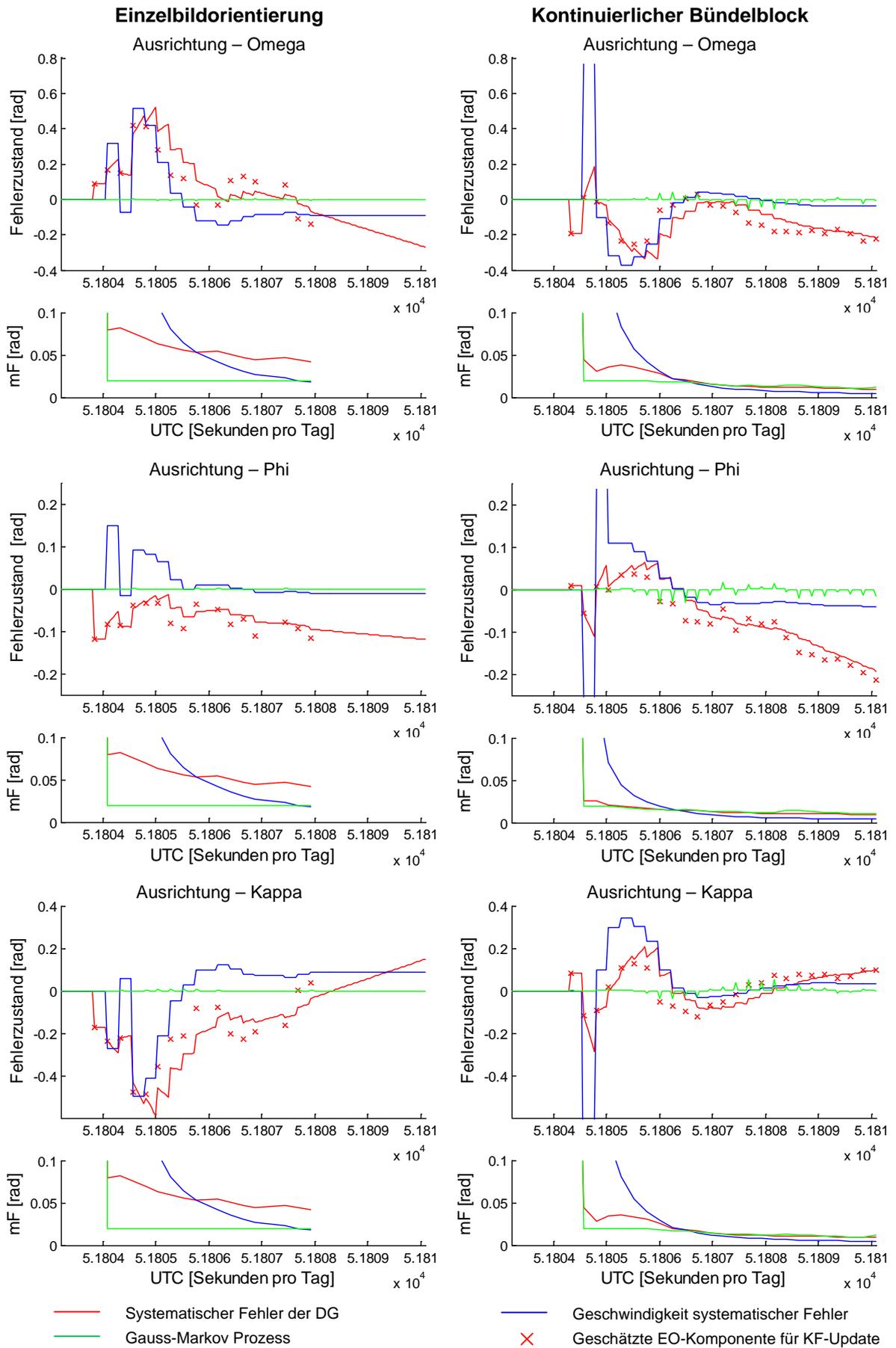


Abbildung 5-8: Äussere Orientierung Testsequenz Ballenberg (Eulerwinkel)

Für statische Aufnahmeconfigurationen bietet damit die Einzelbildorientierung den Vorteil, dass für eine unabhängige Bildorientierungsbestimmung lediglich eine Aufnahme ausreicht, vorausgesetzt die Bild-zu-3D-Stadtmodell-Zuordnung erlaubt die erfolgreiche automatische Generierung von Bildbeobachtungen in allgemeiner Lage.

5.3 Beurteilung und Erfahrungen

Die Funktionstüchtigkeit des in dieser Arbeit vorgeschlagenen zweistufigen Ansatzes zur Georegistrierung mobil erfasster Videosequenzen mit Hilfe digitaler 3D-Landschaftsmodelle konnte nachgewiesen werden. Dabei hängt die erreichbare Video-Georeferenzierungsgenauigkeit der ersten Registrierungsstufe im Wesentlichen von der Qualität der auf der Aufnahmeplattform eingesetzten Navigationssensoren sowie von der realisierten Synchronisationsgenauigkeit zwischen den aufgezeichneten Videodaten und den Navigationsdaten ab. Die eingeführte, auf der direkten Georeferenzierung basierte, Echtzeit-Videoverarbeitung lässt sich ohne Mehraufwand in Kombination mit grösseren mobilen Aufnahmeplattformen und Navigationssensoren hoher Qualität (geodätische GNSS-Empfänger in Kombination mit *tactical grade* INS) einsetzen. In einer solchen Konfiguration ist eine Georegistrierungsgenauigkeit in der Grössenordnung von wenigen Dezimetern möglich, für welche sich eine zusätzliche Verarbeitung mit den Algorithmen der zweiten Registrierungsstufe für die meisten Anwendungsfälle erübrigt. Die zweite Registrierungsstufe ermöglicht eine Verbesserung der Georegistrierungsgenauigkeit um den Faktor drei gegenüber der direkten Georeferenzierungslösung basierend auf Navigationsdaten geringer Qualität. Die integrierte Georeferenzierung von Videodaten basierend auf Navigationsdaten geringer Qualität erfordert eine dynamische Modellierung der enthaltenen systematischen Fehler im Vergleich zu den in der Photogrammetrie etablierten Verfahren. Herauszustreichen ist, dass der für die Realisierung des integrierten Ansatzes implementierte automatische Bild-zu-Modell-Zuordnungsprozess für diverse weitere Anwendungen eingesetzt werden kann, welche die Registrierung beliebig erfasster Bilddaten auf vorliegende 3D-Vektordaten zum Ziel haben. Voraussetzung dafür ist eine Näherungslösung für die Position und Ausrichtung der Bilddaten bezüglich der 3D-Vektordaten.

5.3.1 Unzulänglichkeiten und Schwächen

Die implementierten Algorithmen wurden mit Hilfe von realen Testsequenzen, welche im Aussenraum unter praxisnahen Bedingungen aufgezeichnet wurden, untersucht. Zu erwähnen bleibt, dass die ausgewählten Testgebiete einfache Gebäudestrukturen mit einheitlicher Dach- und Fassadenfarbgebung enthalten. Damit eignen sich diese Testgebiete um die grundsätzliche Tauglichkeit des integrierten Ansatzes zu verifizieren. Jedoch können dicht bebaute städtische Umgebungen komplexere Baustrukturen aufweisen, an welchen die Tauglichkeit der implementierten Konzepte

noch zu prüfen wäre. Für solche Regionen ist die in der Bild-zu-Modell-Zuordnung implementierte Merkmalsextraktion und -interpretation noch ungenügend. Weiter bedingen speziell diese Gebiete eine adäquate Referenzgeometrie, welche grob der Detaillierungsstufe LOD 2 (LOD - engl. *Level of Detail*), wie diese von CityGML (Gröger et al., 2008) spezifiziert wird, entspricht.

Weitere in dieser Arbeit wenig untersuchte und berücksichtigte Effekte sind die wechselnde Belichtung und miterfasste Schlagschatten in den unter realen Einsatzbedingungen aufgezeichneten Videodaten. Im Speziellen bei Einsätzen mit gerichteter Sonneneinstrahlung können diese Effekte für eine robuste integrierte Georeferenzierung nicht vernachlässigt werden und müssen bei der Merkmalsextraktion und Modellbildung im Zuordnungsprozess berücksichtigt werden. Einige Testauswertungen zeigen jedoch, dass das zugrunde gelegte relationale Zuordnungsverfahren, aufgrund der eingeführten Beziehungen zwischen den Primitiven in der Lage ist, extrahierte Schlagschattenkanten von realen Gebäudekanten zu unterscheiden.

Im Weiteren zeigen die Untersuchungen, dass es sehr schwierig ist, die grosse Anzahl Steuerparameter (vgl. Abschnitt 4.3.3) auf die unter unterschiedlichen Bedingungen aufgezeichneten Sequenzen so anzupassen, dass die integrierte Lösung vernünftige Resultate liefert. Aufgrund dieser Tatsache ist eine direkte Umsetzung der implementierten zweiten Registrierungsstufe noch nicht praxisreif. Speziell für Videosequenzen, welche unter realen Bedingungen im Aussenraum aufgezeichnet werden, sind die Parameter für eine erfolgreiche Extraktion der Modellbeschreibung als Basis für den Zuordnungsprozess schwierig festzulegen. Eine robuste Lösung erfordert daher eine Verbesserung dieser Algorithmen. Weitere relevante Parameter, welche die Stochastik der Zustandschätzung beschreiben, könnten unter guter Kenntnis der verwendeten Navigationssensorik und Aufnahmebedingungen automatisch festgelegt werden.

5.3.2 Echtzeitfähigkeit

Die Datenverarbeitung der ersten Registrierungsstufe ist vollständig Echtzeitfähig und steht in der UAVision Softwarelösung zur Verfügung (vgl. Abschnitt 4.2). Die zweite Registrierungsstufe ist lediglich im VGR-Prototyp (vgl. Abschnitt 4.3) implementiert. Die Echtzeitfähigkeit der vorgeschlagenen Algorithmik ist noch nicht demonstriert. Kritischer Algorithmus für eine Echtzeitunterstützung der vorgeschlagenen Lösung ist der implementierte Bild-zu-Modell-Zuordnungsprozess. Insbesondere ist der Berechnungsaufwand der A* Algorithmus-basierten Optimierung der angesetzten Evaluationsfunktion (46) nicht linear zu der Anzahl der involvierten Primitiven in der Modellbeschreibung. Vielmehr zeigen empirische Beobachtungen, dass der Optimierungsaufwand stark von der Genauigkeit der verfügbaren initialen äusseren Orientierungsparameter abhängt. Wie in den Untersuchungen in Abschnitt 3.3.2.1 zum geeigneten Bild-zu-Modell-Zuordnungsverfahren dargestellt, ist lediglich die relationale Zuordnung bei ungenauen initialen Näherungswerten im Stande, eine zuverlässige Lösung zu bestimmen. Damit bietet sich für die Problemlösung der folgende Kompromiss

an: Die Bild-zu-Modell-Zuordnung müsste in einen eigenen *Thread* ausgelagert werden, was die Möglichkeit eröffnet, dass die Systemzustandsbestimmung rückwirkend für eine Epoche geschätzt werden kann. Die direkte Georeferenzierungslösung wird weiter in Echtzeit mit der aktuellsten Fehlerzustandsschätzung korrigiert. Nach der erfolgreichen Schätzung des aktuellen Fehlerzustandes können für die Initialisierung des Zuordnungsprozesses verbesserte Näherungswerte verwendet werden, mit welchen der Optimierungsaufwand deutlich reduziert werden kann. Eine weitere Möglichkeit besteht darin, dass nach einer erfolgreichen Initialisierung des Systemfehlerzustandes von der relationalen Zuordnung beispielsweise auf die diskutierte merkmalsbasierte Zuordnung (vgl. Abschnitt 3.3.2.1) oder auf den von Pilu (1997) vorgeschlagenen Ansatz (vgl. Abschnitt 2.2.2.1) gewechselt wird. Die weiteren im integrierten Ansatz involvierten Algorithmen für die unabhängige Bildorientierungsbestimmung und die kontinuierliche Fehlerzustandsschätzung weisen einen spezifizierbaren Berechnungsaufwand auf. Der Aufwand der kontinuierlichen Bündelblockausgleichung hängt im Wesentlichen von der Anzahl im Bündel mitverarbeiteter Schlüsselbilder und der aktuellen Anzahl Verknüpfungspunkte ab. Die Untersuchungen von Engels et al. (2006) zeigen, dass eine Bündelblockberechnung in Echtzeit möglich ist. Der implementierte Algorithmus kann weiter beschleunigt werden, indem direkt die Normalgleichungsmatrix, wie aus der Online-Triangulation (vgl. Abschnitt 2.3.1.3) bekannt, aktualisiert wird. Weitergehend lässt sich die kontinuierliche Fehlerzustandsschätzung in denjenigen Situationen beschleunigen, in denen sich das Filter im *stady state* Status befindet und damit der Kalman-Filter Gain nicht für jede Zustandsaktualisierung neu berechnet werden muss.

6 Fazit und Ausblick

6.1 Zusammenfassung

Ziel der vorliegenden Arbeit war die Echtzeit-Georegistrierung mittels Mini- und Mikrodrohnen erfasster Videodaten auf ein bestehendes digitales 3D-Stadtmodell unter Zuhilfenahme der verfügbaren Navigationsdaten aus dem Fluglagekontroller und der Stadtmodelldaten. Aus Kosten- und Gewichtsgründen kommen auf solchen Trägersystemen meist Navigationssensoren geringer Qualität zum Einsatz. Für die Lösung dieser Problemstellung wurde ein zweistufiger Video-Georeferenzierungsansatz entwickelt und implementiert. Die erste Registrierungsstufe basiert auf dem Ansatz der direkten Georeferenzierung und nutzt für die Georegistrierung der Videodaten direkt die verfügbaren Fluglagedaten. Der entwickelte Ansatz ermöglicht die Echtzeit Video-Georeferenzierung und berücksichtigt speziell die Eigenschaften und Aufnahme-konfigurationen unbemannter Flugplattformen. Die zweite Registrierungsstufe integriert zusätzlich das digitale 3D-Stadtmodell zur Verbesserung der Georegistrierungsgenauigkeit und demonstriert zudem einen Lösungsansatz für die INS/GNSS/Monokameragestützte Datenintegration. Die Umsetzung der integrierten Lösung erforderte die Entwicklung und Implementierung folgender Teilkomponenten. Als erstes wurde ein robustes automatisches Zuordnungsverfahren evaluiert und entwickelt, welches die erfassten Bilddaten in Beziehung zu der bestehenden 3D-Referenzgeometrie setzt. Basierend auf dieser Zuordnung ermöglichen anschliessend die beiden implementierten Ansätze 'Rückwärtsschnitt' und 'kontinuierlicher Bündelblock' die unabhängige Schätzung der Orientierungsparameter ausgewählter Video-Einzelbilder. Abschliessend können mit dem vorgeschlagenen komplementären Kalman-Filter mit Hilfe der unabhängig bestimmten Orientierungsparameter der aktuelle systematische Fehler der direkten Georeferenzierungslösung geschätzt und berücksichtigt werden. Die erste Registrierungsstufe wurde sowohl innerhalb der Softwarelösung UAVision wie auch im VGR-Prototyp implementiert. Das vorgeschlagene Konzept der zweiten Registrierungsstufe wurde zwecks Genauigkeitsuntersuchungen und Validierung lediglich im VGR-Prototyp umgesetzt. Die UAVision-Implementierung unterstützt im Gegensatz zum VGR-Prototyp die Echtzeitdatenverarbeitung.

Der vorgeschlagene Ansatz unterscheidet sich von diversen INS/GNSS/Monokamera-Integrationslösungen im Wesentlichen dadurch, dass er eine Zuordnung zwischen den erfassten Bilddaten und einer bestehenden Geodatengrundlage realisiert. Damit resultieren absolute anstatt relative Orientierungsupdates für die Trägerplattform aus der Monokamerakonfiguration als Input für die Datenintegration. Verglichen mit der in der Photogrammetrie etablierten direkten und integrierten Bildorientierung berücksichtigt der vorgeschlagene Ansatz die Fehltausrichtung und Hebelarmkomponente kombiniert mit den korrelierten systematischen Fehlern der Navigationslösung in einem kontinuierlich geschätzten Fehlerterm.

Für die Validierung und Genauigkeitsuntersuchungen der implementierten Konzepte wurden unterschiedliche mittels Mikro- und Minidrohnen system aufgezeichnete Videosequenzen verwendet. Aufgrund der erfassten Testsequenzen konnte die grundsätzliche Funktionstüchtigkeit der beiden Registrierungsstufen nachgewiesen werden. Die durchgeführten Genauigkeitsuntersuchungen zeigten, dass der integrierte Ansatz eine Steigerung der Georegistrierungsgenauigkeit durchschnittlich um den Faktor drei verglichen zur direkten Georeferenzierungslösung ermöglicht. Abhängig von der Flugstabilität der verwendeten Aufnahmeplattform, der Qualität der Navigationssensoren sowie der verfügbaren Fluglagedaten-Updatefrequenz konnten Georegistrierungsgenauigkeiten für Minidrohnen im Bereich von 0.5m und für Mikrodrohnen von 1m erzielt werden. Weiter zeigten zusätzliche Untersuchungen, dass gerade bei starker Windbeeinflussung eine hohe Fluglagedaten-Updatefrequenz (> 25 Hz) für eine erfolgreiche integrierte Lösung notwendig ist. Liegen die Fluglagedaten in einer zu geringen Frequenz vor, konnten die hochfrequent auftretenden Plattformbewegungen – verursacht beispielsweise durch Windböen – nicht erfasst und berücksichtigt werden. Weiter konnte gezeigt werden, dass die unabhängige äussere Orientierungsbestimmung die Qualität des geschätzten aktuellen Fehlerzustandes beeinflusst. Entsprechend aufgezeichnete Testsequenzen belegen, dass der kontinuierliche bündelblockbasierte Ansatz für ein dynamisch aufgezeichnetes Video im Vergleich zum rückwärtsschnittbasierten Ansatz die besseren Resultate liefert. Im Gegensatz dazu liefert der rückwärtsschnittbasierte Ansatz für nahezu statisch aufgezeichnete Videos deutlich bessere Georegistrierungsgenauigkeiten.

6.2 Empfehlungen

Der vorgestellte integrierte Ansatz erlaubt die Schätzung und Kompensation verbleibender Restsystematiken zwischen der vorhandenen direkten Georeferenzierungslösung und vorliegenden 3D-Referenzdaten und eignet sich damit optimal für die Umsetzung eines AR-Szenarios. Das für die integrierte Georeferenzierung vorgeschlagene Bild-zu-Modell-Zuordnungsverfahren kann zusätzlich für die Lösung diverser Problemstellungen unabhängig von der in dieser Arbeit aufgezeigten Anwendung genutzt werden. Der vorgeschlagene Algorithmus erlaubt die automatische Zuordnung beliebig erfasster Bilddaten auf bestehende 3D-Stadtmodellgeometrien ausgehend von einer mit Hilfe kostengünstiger Navigationssensoren bestimmten Näherungslösung für Aufnahmeposition und -ausrichtung. Die erfolgreiche Bild-zu-Modell-Zuordnung bildet die Basis für den in dieser Arbeit vorgeschlagenen integrierten Georeferenzierungsansatz. Jedoch zeigen erste Erfahrungen aus den ausgewerteten Testsequenzen, dass gerade die zuverlässige Zuordnung aufgrund stark wechselnder Aufnahmebedingungen, eine sehr schwierige Aufgabe ist. Herausforderung dabei ist im Speziellen die automatische robuste Extraktion der für die relationale Zuordnung benötigten Modellbeschreibung aus den Bilddaten. Die dafür

implementierten Algorithmen sind noch zu stark abhängig von verschiedenen Schwellwerten, die aufgrund der vorherrschenden Aufnahmebedingungen variieren und schwierig automatisch festzulegen sind. Eine weitere Verbesserung des Zuordnungsproblems könnte durch die zusätzliche Verwendung von zur Verfügung stehenden Sachdaten im 3D-Stadtmodell im relationalen Zuordnungsprozess erzielt werden. Abschliessend sei aus aktueller Sicht bemerkt, eine optimale Video-Georeferenzierungsgenauigkeit kann, bei ausreichender Nutzlast der Trägerplattform zum einen und finanziellen Mitteln zum anderen, durch den Einsatz von INS/GNSS-Sensoren hoher Qualität zuverlässiger realisiert werden als durch die hier vorgeschlagene softwarebasierte integrierte Lösung in Kombination mit Navigationssensoren geringer Qualität.

6.3 Anwendungsmöglichkeiten

Die vorgeschlagenen Konzepte und implementierten Algorithmen können die Umsetzung unterschiedlicher Anwendungsszenarien unterstützen. Die vorgestellte zweistufige Video-Georeferenzierungslösung lässt sich sehr einfach integrieren in Kombination mit unterschiedlichen mobilen Aufnahmeträgerplattformen. Die zweistufige Architektur erlaubt zudem je nach anwendungsspezifischen Anforderungen die flexible Wahl zwischen direkter und integrierter Georeferenzierungslösung. Zudem ermöglicht die gewählte Architektur die einfache Integration und Nutzung von Navigationssensoren unterschiedlicher Güte und von verschiedenen Herstellern.

Speziell interessant ist die vorgeschlagene Bild-zu-Modell-Zuordnung für die Umsetzung verschiedener *Location Based Services* (LBS) Anwendungen im städtischen Umfeld. Viele heute auf dem Markt erhältliche mobile Geräte enthalten nebst einer Kamera eine einfache Navigationssensorik zur näherungsweise Bestimmung der Position und Ausrichtung im Raum. Basierend auf diesen Informationen und dem Resultat aus der Bild-zu-Modell-Zuordnung gestatten nun die Algorithmen zur unabhängigen Bestimmung der äusseren Orientierung eine automatische präzise Ausrichtung der mobilen Geräte auf bestehende 3D-Stadtmodellgeometrien.

Die integrierte Lösung aus INS/GNSS/Monokamera steht nicht nur für die Georeferenzierung mobil erfasster Videodaten zur Verfügung, sondern könnte auch für die optimale Schätzung der Positionierung und Ausrichtung innerhalb der Navigationsrechnung verwendet werden. Speziell die realisierte unabhängige monokamerabasierte Orientierungsbestimmung könnte das INS während GNSS-Signalausfällen stützen. Weiter ist der vorgeschlagene Ansatz für SLAM-Anwendungen interessant, in welchen die simultan kartierten Objekte und die bestimmte Aufnahmetrajektorie in einen bestehenden Geodatenbestand bezüglich eines übergeordneten Referenzkoordinatensystems integriert werden müssen. Die eingeführte bündelblockbasierte Bildorientierung könnte weiterführend auch für die Umsetzung einer SfM-

Lösung Verwendung finden. Eine erste automatische Rekonstruktion der Struktur resultiert aus den während der Verarbeitung bestimmten Verknüpfungspunkten.

6.4 Ausblick

Eine vollständige Integration und Umsetzung der zweiten Registrierungsstufe in einer operativen Lösung bedingt in erster Linie die Verbesserung und Weiterentwicklung der Algorithmen zur automatischen Extraktion der Modellbeschreibung im Bild-zu-Modell-Zuordnungsprozess. Weiter ist die implementierte integrierte Lösung nicht für eine Echtzeitverarbeitung ausgelegt und müsste entsprechend wie in Abschnitt 5.3.2 bereits angesprochen weiterentwickelt werden. Viele in dieser Arbeit implementierte Algorithmen können hinsichtlich Performance und Robustheit stark verbessert werden. Gerade in diesem Zusammenhang müsste die kontinuierliche bündelblock- und rückwärtsschnittbasierte äussere Orientierungsbestimmung durch einen robusten Schätzer ersetzt werden. Hier bieten sich beispielsweise die in der Photogrammetrie gebräuchlichen Schätzer von Klein und Förstner (1984) oder Kruck (1995) an. Der Berechnungsaufwand der kontinuierlichen Bündelblockausgleichung liesse sich weiter mit Hilfe von Verfahren der Online-Triangulation (Grün, 1982) zwischen nachfolgenden Bündelberechnungen reduzieren.

Die integrierte Lösung im implementierten VGR-Prototyp bezieht die bestehende 3D-Statdmodellgeomtrie direkt aus einer für Testzwecke umgesetzten Geodatenbasis im Arbeitsspeicher. Die direkte Anbindung an den virtuellen Globus i3D, aus welchem die benötigten 3D-Referenzgeometriedaten für den Bild-zu-Modell-Zuordnungsprozess hoch performant zur Verfügung gestellt werden sollen, bleibt noch zu realisieren und auf ihre Tauglichkeit hin zu verifizieren. Die Nutzung des virtuellen Globus für die Referenzgeometrieextraktion aus dem bestehenden Geodatenbestand, würde auf der einen Seite die Performance des Zuordnungsprozesses erhöhen und auf der anderen Seite könnte die Extraktion der Modellbeschreibung sowie der Zuordnungsprozess aufgrund verfügbarer Sachdaten verbessert werden.

Literatur

- Abdel-Aziz, Y. I. und Karara, H. M. (1971). Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. Symposium on Close-Range Photogrammetry, Falls Church, VA, American Society of Photogrammetry.
- Albertz, J. und Wiggenhagen, M. (2009). Taschenbuch zur Photogrammetrie und Fernerkundung: Guide for Photogrammetry and Remote Sensing. 5. Auflage, Heidelberg, Wichmann.
- Bailey, T. und Durrant-Whyte, H. F. (2006). Simultaneous Localisation and Mapping (SLAM): Part II State of the Art. IEEE Robotics & Automation Magazine, 13(3) pp. 108-117.
- Ballard, D. H. (1987). Generalizing the hough transform to detect arbitrary shapes. Pattern Recognition, 13(2) pp. 111-122.
- Baltsavias, E. P. (1991). Multiphoto geometrically constrained matching. 49, Zürich, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich.
- Barrow, H. G., Tenenbaum, J. M., Bolles, R. C. und Wolf, H. C. (1977). Parametric correspondence and chamfer matching: Two new techniques for image matching. In Proc. 5th international joint conference on artificial intelligence, pp. 659-663. Cambridge, Mass.
- Bäumker, M. (2007). Kalibration und Umrechnung von INS- und photogrammetrischen Winkeln für beliebige gegenseitige Anordnungen. 14. Internationale Geodätische Woche 2007, Obergurgl, 11.-17. Februar.
- Bayoud, F. A. (2006). Development of a robotic mobile mapping system by vision-aided inertial navigation a geomatics approach. Geodätische-geophysikalische Arbeiten in der Schweiz, 71, Zürich, Schweizerische Geodätische Kommission.
- Bento, M. (2008). Unmanned Aerial Vehicles: An Overview. InsideGNSS, January/February pp. 54-61.
- Beveridge, J. R. und Riseman, E. M. (1997). How Easy is Matching 2D Line Models Using Local Search?. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(6) pp. 564-579.
- Borgefors, G. (1986). Distance transformations in digital images. Computer Vision, Graphics and Image Processing, 34(3) pp. 344-371.
- Borgefors, G. (1988). Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm. IEEE Transaction on Pattern Analysis and Machine Intelligence, 10(6) pp. 849-865.
- Boyer, K. L. und Kak, A. C. (1986). Symbolic Stereo from Structural Descriptions. School of Electrical Engineering, Purdue University, Technical Report TR-EE 86-12.
- Boyer, K. L. und Kak, A. C. (1988). Structural stereopsis for 3-D vision. IEEE Transactions on Pattern Analysis and Machine Intelligence, 10(2) pp. 144-166.
- Brown, D. C. (1971). Close-Range Camera Calibration. Photogrammetric Engineering, 37(8) pp. 855-866.

- Brown, R. G. und Hwang, P. Y. C. (1997). Introduction to random signals and applied Kalman filtering with Matlab exercises and solutions. 3th Auflage, New York, John Wiley.
- Cannata, R. W., Shah, M., Blask, S. G. und Van Workum, J. A. (2000). Autonomous Video Registration Using Sensor Model Parameter Adjustment. In Proc. 29th Applied Imagery Pattern Recognition Workshop, pp. 215 - 222. AIPR'00, Washington, DC, USA.
- Canny, J. F. (1983). Finding Edges and Lines in Images. Massachusetts Institute of Technology, Artificial Intelligence Laboratory, Technical Report: AITR-720
- Cho, W. (1996). Relational Matching for Automatic Orientation. In Proc. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 31(Part B3), pp. 111-119. XVIII ISPRS Congress, Vienna.
- Christen, M. und Nebiker, S. (2010). Neue Entwicklungen im Bereich Virtueller Globen am Beispiel der i3D-Technologie. In Proc. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V., 19, pp. 325-333. Dreiländertagung DGPF - OVG - SGPF, Wien.
- Cocard, M. (1995). High precision GPS processing in kinematic mode. Geodätische-geophysikalische Arbeiten in der Schweiz, 52, Zürich, Schweizerische Geodätische Kommission.
- Cornelis, K., Pollefeys, M., Vergauwen, M. und Van Gool, L. (2001). Augmented Reality using Uncalibrated Video Sequences. In: Lecture Notes in Computer Science. Springer. 2018 pp. 144-160.
- Craig, J. J. (1986). Introduction to robotics mechanics & control. Reading - Mass., Addison-Wesley.
- Cramer, M. (2001a). Genauigkeitsuntersuchung zur GPS/INS-Integration in der Aerophotogrammetrie. Dissertation, Stuttgart, Universität Stuttgart.
- Cramer, M. (2001b). Performance of GPS/Inertial Solutions in Photogrammetry. In Proc. Photogrammetric Week '01', pp. 49-62. Photogrammetric Week '01', Stuttgart, Wichmann, Heidelberg.
- Davison, A. J. (2003). Real-Time Simultaneous Localisation and Mapping with a Single Camera. In Proc. IEEE International Conference on Computer Vision, pp. 1403–1410. Nice.
- Dobbert, T. (2005). Matchmoving: The invisible art of camera tracking. Alameda, California.
- Duda, R. O. und Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. Communications of the ACM, 15(1) pp. 11-15.
- Durrant-Whyte, H. F. und Bailey, T. (2006). Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. Robotics & Automation Magazine, IEEE, 13(2) pp. 99-110.
- Durrant-Whyte, H. F., Rey, D. und Nebot, E. (1996). Localisation of automatic guided vehicles. In Proc. 7th International Symposium Robotics Research, pp. 613-625. ISRR'95, Springer Verlag.

- Ebner, H. (1976). Self-calibrating bundle adjustment. In Proc. International Archives of the Photogrammetry and Remote Sensing, 13. XIII ISPRS Congress, Helsinki.
- Eck, C. (2001). Navigation algorithms with applications to unmanned helicopters. PhD Thesis, Zürich, ETH Zürich.
- Eisenbeiss, H. (2009). UAV photogrammetry. 105, Zürich, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich.
- El-Sheimy, N. (2009). Emerging MEMS IMU and Its Impact on Mapping Applications. In Proc. Photogrammetric Week '09, pp. 203-216. Photogrammetric Week '09, Stuttgart, Wichmann.
- Ellum, C. M. und El-Sheimy, N. (2006). New Strategies for intergrating photogrammetric and GNSS data. In Proc. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(Part 5), pp. 103-108. ISPRS Commission V Symposium 'Image Engineering and Vision Metrology', Dresden.
- Engels, C., Stewénius, H. und Nistér, D. (2006). Bundle Adjustment Rules. Photogrammetric Computer Vision (PCV), Bonn, D.
- Eugster, H., Flückiger, K. und Nebiker, S. (2010). Drohnenbasierte Umweltbeobachtung und Kartierung basierend auf einem Virtuellen Globus. In. Dreiländertagung DGPF - OVG - SGPF, Wien, A.
- Eugster, H. und Nebiker, S. (2007). Geo-registration of Video Sequences captured from mini UAVs – Approaches and Accuracy Assesment. 5th International Symposium on Mobile Mapping Technology, Padua, I.
- Eugster, H. und Nebiker, S. (2008). UAV-Based Augmented Monitoring - Real-Time Georeferencing and Integration of Video Imagery with Virtual Globes. In Proc. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(Part B1), pp. 1229-1236. XXI ISPRS Congress, Beijing.
- Faig, W. (1975). Calibration of Close-Range Photogrammetric Systems: Mathematical Formulation. Photogrammetric Engineering and remote sensing, 41(12) pp. 1479-1486.
- Farley, M. G. und Chapman, M. D. (2008). Monocular SLAM Alternative Navigation for GPS-Denied Areas. GPS World, (September) pp. 42-49.
- Faugeras, O. D. (1993). Three-dimensional computer vision a geometric viewpoint. Cambridge, Mass., MIT Press.
- Favey, E. (2001). Investigation and Improvement of airborne laser scanning technique for monitoring surface elevation changes of glaciers. 72, Zürich, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich.
- Fischler, M. und Bolles, R. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Communications of the ACM, 24(6) pp. 381-395.
- Förstner, W. und Gülch, E. (1987). A Fast Operator for Detection and Precise Location of Distict Point, Corners and Centres of Circular Features. ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data, Interlaken, CH.

- Gelb, A. (1986). Applied optimal estimation. 9th Auflage, Cambridge, Mass., MIT Press.
- Grafarend, E., Lohse, P. und Schaffrin, B. (1989). Dreidimensionaler Rückwärtsschnitt. Zeitschrift für Geodäsie, Geoinformation und Landmanagement (zfv), (114) pp. 61-67,127-137,172-175,225-234,278-287.
- Grenzdörffer, G. und Engel, A. (2008). Eine vergleichende Untersuchung von zwei Micro-UAV's - Perspektiven für die aktuelle und kostengünstige Geoinformationsgewinnung. GIS Zeitschrift für Geoinformatik, 1/2008 pp. 17-25.
- Gröger, G., Kolbe, T., Czerwinski, A. und Nagel, C. (2008). OpenGIS City Geographics Markup Language (CityGML). Open Geospatial Consortium Inc., OGC 08-007r1.
- Grün, A. (1982). An Optimum Algorithm for On-Line Triangulation. Symposium of the Commission III of the ISPRS "Mathematical Models, Accuracy Aspects and Quality Control", Helsinki.
- Grün, A. (1985). Algorithmic Aspects in On-Line Triangulation. Photogrammetric Engineering and Remote Sensing, 51(4) pp. 419-436.
- Haala, N. (2005). Multi-Sensor-Photogrammetrie. Habilitation, Stuttgart, Universität Stuttgart.
- Haala, N., Böhm, J. und Kada, M. (2002). Processing of 3D building models for location aware applications. In Proc. International Archives on Photogrammetry and Remote Sensing pp. 138-143. ISPRS Commission III Symposium, Graz, A.
- Harris, C. und Stephens, M. (1988). A Combined Corner and Edge Detector. In Proc. The Fourth Alvey Vision Conference, pp. 147-151.
- Hartley, R. und Zisserman, A. (2003). Multiple view geometry in computer vision. 2nd Auflage, Cambridge, Cambridge University Press.
- Heipke, C. (1996). Overview of Image Matching Techniques. OEEPE Workshop on the Application of Digital Photogrammetric Workstations, Lausanne, CH.
- Heipke, C., Jacobsen, K. und Wegmann, H. (2002). Analysis of the results of the OEEPE test "Integrated Sensor Orientation". In: Heipke, C., Jacobsen, K. and Wegmann, H. ed. Integrated Sensor Orientation Test Report and Workshop Proceedings. European Organization for Experimental Photogrammetric Research. 43 pp. 31-49.
- Hirschmüller, H. (2008). Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2) pp. 328-341.
- Hummel, R. A. und Zucker, S. W. (1987). On the Foundation of Relaxation Labeling Processes. In: Fischler, M. and Firschein, O. ed. Readings in computer vision: issues, problems, principles, and paradigms. San Francisco, CA, USA, Morgan Kaufmann Publishers Inc. pp. 585-605.
- Jähne, B. (2005). Digitale Bildverarbeitung. 6. Auflage, Berlin, Springer.
- Jekeli, C. (2001). Inertial navigation systems with geodetic applications. Berlin, Walter de Gruyter.

- Julier, S. J. und Uhlmann, J. K. (1997). A New Extension of the Kalman Filter to Nonlinear Systems. In Proc. Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, pp. 182-193. Orlando, FL.
- Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME – Journal of Basic Engineering, 82(Serie D) pp. 35-45.
- Klein, H. und Förstner, W. (1984). Realization of Automatic Error Detection in the Block Adjustment Program PAT-M43 Using Robust Estimators. In Proc. International Archives of Photogrammetry and Remote Sensing, 25(Part A3a). XV ISPRS Congress, Rio de Janeiro.
- Kozempel, K. und Reulke, R. (2009). Camera orientation based on matching road networks. In Proc. Image and Vision Computing New Zealand, 2009, pp. 237 - 242 IVCNZ '09, Wellington
- Kraus, K., Jansa, J. und Kager, H. (1996). Photogrammetrie Band 2 - Verfeinerte Methoden und Anwendungen. 3. Auflage, Bonn, Dümmler.
- Kraus, K. und Waldhäusl, P. (1997). Photogrammetrie Band 1 - Grundlagen und Standardverfahren. 6. Auflage, Bonn, Dümmler.
- Kruck, E. (1995). Balanced Least Squares Adjustment for Relative Orientation. In: Grün, A. and Kahmen, H. ed. Optical 3-D Measurement Techniques III. Wichmann.
- Kumar, R., Sawhney, H. S., Asmuth, J. C., Pope, A. und Hsu, S. (1998). Registration of video to geo-referenced imagery. In Proc. 14th International Conference on Pattern Recognition, 2. ICPR '98, Brisbane.
- Leick, A. (2004). GPS satellite surveying. 3rd Auflage, Hoboken, NJ., Wiley.
- Leonard, J. J. und Durrant-Whyte, H. F. (1991). Simultaneous Map Building and Localization for an Autonomous Mobile Robot. In Proc. IEEE/RSJ International Workshop on Intelligent Robots and Systems, 3 pp. 1442 - 1447. IROS '91, Osaka, Japan.
- Levenberg, K. (1944). A method for the solution for certain nonlinear problems in least-squares. The Quarterly Journal of Mechanics and Applied Mathematics, 2 pp. 164-168.
- Lowe, D. (1999). Object Recognition from Local Scale-Invariant Features. In Proc. 7th International Conference on Computer Vision, 2, pp. 1150-1157. Corfu, Greece.
- Luber, A. und Reulke, R. (2009). Validation of a polynomial camera model for generic cameras: Calibration of catadioptric, fish-eye, short- and long focal length cameras. In Proc. Image and Vision Computing New Zealand, 2009, pp. 179 - 184 IVCNZ '09, Wellington.
- Luhmann, T. (2003). Nahbereichsphotogrammetrie Grundlagen, Methoden und Anwendungen. 2. Auflage, Heidelberg, Wichmann.
- Luhmann, T. (2010). Erweiterte Verfahren zur geometrischen Kamerakalibrierung in der Nahbereichsphotogrammetrie. C 645, Deutsche Geodätische Kommission (DGK)

- Madsen, K., Nielsen, H. B. und Tingleff, O. (1999). *Methods for Non-linear Least Squares Problems*. Informatics and Mathematical Modelling, Technical University of Denmark.
- Mäntylä, M. (1984). *An Introduction to Solid Modeling*. Laboratory of Information Processing Science, Helsinki University of Technology, Report-HTKK-TKO.
- Marquardt, D. W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2) pp. 431-441
- Matai, B., Wides, R. P., Hsu, S., Kumar, R., Samarasekera, S., Sawhney, H. S. und Hanna, K. (2003). Robust Video Georegistration in the Presence of Significant Appearance Changes. In: Shah, M., Kumar, R. ed. *Video registration*. Boston, Mass., Kluwer Academic pp. 180-221.
- Mautz, R. (2001). *Zur Lösung nichtlinearer Ausgleichsprobleme bei der Bestimmung von Frequenzen in Zeitreihen*. Dissertation, Berlin, Technische Universität Berlin.
- McIntosh, J. H. und Mutch, K. M. (1988). Matching Straight Lines. *Computer Vision, Graphics, and Image Processing*, 43(3) pp. 386-408.
- Mittermayer, E. (1972). Zur Ausgleichung freier Netze. *Zeitschrift für Geodäsie, Geoinformation und Landmanagement (zfv)*, 97(11) pp. 481-489.
- Montiel, J. M. M., Civera, J. und Davison, A. J. (2006). *Unified Inverse Depth Parametrization for Monocular SLAM*. Robotics: Science and Systems Philadelphia.
- Moravec, H. (1977). Towards automatic visual obstacle avoidance. 5th International Joint Conference on Artificial Intelligence, MIT, Cambridge, Mass.
- Mostafa, M. und Hutton, J. (2005). 10 Years of Direct Georeferencing for Airborne Photogrammetry. *GIS*, 11/2005.
- Müller, F. (1991). *Photogrammetrische Punktbestimmung mit Bilddaten digitaler Dreizeilenkameras*. DGK Reihe C, 372, München, Deutsche Geodätische Kommission.
- Nebiker, S., Annen, A., Scherrer, M. und Oesch, D. (2008). A Light-weight Multispectral Sensor for Micro UAV – Opportunities for Very High Resolution Airborne Remote Sensing. In *Proc. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(Part B1), pp. 1193-1200. XXI ISPRS Congress, Beijing.
- Nebiker, S., Bleisch, S. und Gülch, E. (2010). Virtual Globes. *GIM International*, 24(7).
- Nelder, J. A. und Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, 7(4) pp. 308-313.
- Pearl, J. (1984). *Heuristics intelligent search strategies for computer problem solving*. Reading - Mass.a.o., Addison-Wesley.
- Perelmuter, A. (1979). Adjustment of free Networks. *Bulletin Géodésique*, 53(4) pp. 291-295.

- Pilu, M. (1997). A Direct Method for Stereo Correspondance based on Singular Value Decomposition. IEEE International Conference of Computer Vision and Pattern Recognition (CVPR), Puerto Rico.
- Pollefeys, M. (1999). Self-calibration and Metric 3D Reconstruction from Uncalibrated Image Sequences. PhD Thesis, Leuven, Katholieke Universiteit Leuven.
- Pollefeys, M., Koch, R., Vergauwen, M. und Van Gool, L. (1999). Hand-held acquisition of 3D models with a video camera. In Proc. 2nd International Conference on 3-D Imaging and Modeling, pp. 14-23 3DIM '99, Ottawa.
- Pollefeys, M. und Van Gool, L. (1999). Stratified Self-Calibration with the Modulus Constraint. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(8) pp. 707-724.
- Press, W. H. (2007). Numerical Recipes: The Art of Scientific Computing. 3rd Auflage, Cambridge, Mass., Cambridge University Press.
- Ressl, C. (2003). Geometry, Constraints and Computation of the Trifocal Tensor. PhD Thesis, Wien, Technische Universität Wien.
- Ries, C. (2004). Ein allgemeiner Ansatz zur Georeferenzierung von multispektralen Flugzeugscanneraufnahmen. Dissertation, Wien, Technische Universität Wien.
- Rodehorst, V. (2004). Photogrammetrische 3D-Rekonstruktion im Nahbereich durch Auto-Kalibrierung mit projektiver Geometrie. Berlin, Wissenschaftlicher Verlag Berlin.
- Sandau, R. und Eckert, A. (1996). The stereo Camera Family WAOS/WAAC for Spaceborne / Airborne Applications. In Proc. International Archives of Photogrammetry and Remote Sensing, 31(Part B1), pp. 170-175.
- Schenk, T. (2004). From point-based to feature-based aerial triangulation. ISPRS Journal of Photogrammetry & Remot Sensing, 58(5-6) pp. 315-329.
- Schmid, H. H. (1977). Ein allgemeiner Ausgleichungs-Algorithmus für die numerische Auswertung in der Photogrammetrie. 22, Zürich, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich.
- Schmidt, U. (2005). Professionelle Videotechnik Analoge und digitale Grundlagen, Filmtechnik, Fernsehtechnik, HDTV, Kameras, Displays, Videorecorder, Produktion und Studioteknik. 4. Auflage, Berlin, Springer.
- Scholze, S. (2002). A probabilistic approach to building roof reconstruction. Konstanz, Hartung-Gorre Verlag.
- Schwarz, K. P. (1995). Integrated airborne navigation systems for photogrammetry. In Proc. Photogrammetric Week '95, pp. 139-153. Photogrammetric Week '95, Herbert Wichmann Verlag, Heidelberg.
- Schwarz, K. P., Chapman, M. A., Cannon, M. E. und Gong, P. (1993). An Integrated INS/GPS Approach to the Georeferencing of Remotely Sensed Data. Photogrammetry Engineering & Remote Sensing, 59(11) pp. 1667-1674.
- Schwarz, K. P., Fraser, C. S. und Gustafson, P. C. (1984). Aerotriangulation without Ground Control. In Proc. International Archives of Photogrammetry and Remote Sensing, 25(Part A1), pp. 237-250. XV ISPRS Congress, Rio de Janeiro.

- Schwermann, R. (1995). Geradengestützte Bildorientierung in der Nahbereichsphotogrammetrie. Ph.D, Aachen, Rheinisch - Westfälische Technische Hochschule Aachen.
- Scott, G. und Longuet-Higgins, H. (1991). An algorithm for associating the features of two patterns. In Proc. Royal Society London, B244, pp. 21-26.
- Sester, M. und Förstner, W. (1989). Object Location based on Uncertain Models. 11. DAGM-Symposium Mustererkennung, Hamburg, Springer-Verlag.
- Shapiro, L. G. und Haralick, R. M. (1981). Structural Description and Inexact Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 3(3) pp. 504-519.
- Shapiro, L. G. und Haralick, R. M. (1987). Relational Matching. Applied Optics, 26(10) pp. 1845-1851.
- Sheikh, Y., Khan, S. und Shah, M. (2005). Feature-Based Georegistration of Aerial Images. In: Stefanidis, A. ed. GeoSensor networks. Boca Raton, Fla., CRC Press pp. 125-147.
- Sheikh, Y., Khan, S., Shah, M. und Cannata, R. (2003). Geodetic Alignment of Aerial Video Frames. In: Shah, M., Kumar, R. ed. Video registration. Boston, Mass., Kluwer Academic pp. 144-179.
- Shi, J. und Tomasi, C. (1994). Good Features to Track. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 593 - 600. CVPR '94, Seattle.
- Shin, E.-H. (2005). Estimation Techniques for Low-Cost Inertial Navigation. PhD, Calgary, University of Calgary.
- Shoemake, K. (1985). Animating Rotation with Quaternion Curves. ACM SIGGRAPH Computer Graphics, 19(3) pp. 245-254.
- Skaloud, J. (1999a). Optimizing Georeferencing of Airborne Survey Systems by INS/DGPS. PhD Thesis, Calgary, University of Calgary.
- Skaloud, J. (1999b). Problems in Direct-Georeferencing by INS/DGPS in the Airborne Environment. ISPRS Workshop 'Direct versus Indirect Methodes of Sensor Orientation', Barcelona.
- Skaloud, J., Cramer, M. und Schwarz, K. P. (1996). Exterior orientation by direct measurement of camera position and orientation. In Proc. International Archives of Photogrammetry and Remote Sensing, 31(Part B3), pp. 125-130. XVIII ISPRS Congress, Vienna.
- Stallmann, D. (2008). DGAP Notes Version 3.2.0, pp. 37-43.
- Steffen, R. und Förstner, W. (2008). On Visual Real Time Mapping for Unmanned Aerial Vehicles. In Proc. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(B3a), pp. 57-62. XXI ISPRS Congress, Beijing.
- Swisstopo (2008). Formeln und Konstanten für die Berechnung der Schweizerischen schiefachsigen Zylinderprojektion und der Transformation zwischen Koordinatensystemen. Bundesamt für Landestopografie swisstopo.

- Thrun, S., Burgard, W. und Fox, D. (2005). Probabilistic robotics. Cambridge, Mass., MIT Press.
- Titterton, D. H. und Weston, J. L. (2004). Strapdown inertial navigation technology. 2nd Auflage, Reston, American Institute of Aeronautics and Astronautics.
- Tomasi, C. und Kanade, T. (1991). Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132.
- Torge, W. (2003). Geodäsie. 2. Auflage, Berlin, W. de Gruyter.
- Triggs, B., McLauchlan, P., Hartley, R. und Fitzgibbon, A. (2000). Bundle Adjustment - A Modern Synthesis. In: Triggs, W., Zisserman, A., Szeliski, R. ed. Lecture Notes In Computer Science. London, Springer Verlag. 1883 pp. 298-375.
- Tsai, R. Y. (1987). A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. IEEE Journal of Robotics and Automation, 3(4).
- Vörsmann, P., Buschmann, M., Krüger, L. und Martin, T. (2007). The MINC System - a highly miniaturized integrated Navigation and Control. The International UAV World Conference, Frankfurt, D.
- Vosselman, G. (1992). Relational matching. Lecture Notes in Computer Science, 628, Berlin etc., Springer Verlag.
- Wägli, A. (2009). Trajectory determination and analysis in sports by satellite and inertial navigation. Geodätische-geophysikalische Arbeiten in der Schweiz, 77, Zürich, Schweizerische Geodätische Kommission.
- Weatherington, D. (2005). Unmanned Aircraft Systems (UAS) Roadmap 2005-2030 (DoD), U. D. o. D.
- Welch, G. und Bishop, G. (2004). An Introduction to the Kalman Filter. Department of Computer Science, University of North Carolina at Chapel Hill.
- Wendel, J., Meister, O., Schlaile, C. und Trommer, G. (2006). An integrated GPS/MEMS-IMU navigation system for an autonomous helicopter. Aerospace Science and Technology, 10(6) pp. 527-533.
- Wewel, F., Scholten, F., Neukum, G. und Albertz, J. (1998). Digitale Luftbildaufnahme mit der HRSC - Ein Schritt in die Zukunft der Photogrammetrie. PFG - Photogrammetrie, Fernerkundung, Geoinformation, 6 pp. 337-348.
- Wicki, F. (1998). Robuste Schätzverfahren für die Parameterschätzung in geodätischen Netzen. 67, Zürich, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich.
- Winkler, S. (2007). Zur Sensordatenfusion für integrierte Navigationssysteme unbemannter Kleinstflugzeuge. Dissertation, Braunschweig, Technische Universität Braunschweig.
- Wolf, H. (1972). Helmerts Lösung zum Problem der freien Netze mit singulärer Normalgleichungsmatrix. Zeitschrift für Geodäsie, Geoinformation und Landmanagement (zfv), 97(5) pp. 189-192.

- Zalmanson, G. H. (2000). Hierarchical Recovery of Exterior Orientation from Parametric and Natural 3-D Curves. In Proc. International Archives of Photogrammetry and Remote Sensing, 33(Part B2). XIX ISPRS Congress, Amsterdam.
- Zanini, M. (1998). Dreidimensionale synthetische Landschaften wissensbasierte dreidimensionale Rekonstruktion und Visualisierung raumbezogener Informationen. 66, Zürich, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich.

Anhang

A Parametrisierung Rotationsmatrizen

Gebräuchliche Parametrisierung in der Photogrammetrie

Drehung zwischen Bildkoordinaten- und übergeordnetem Referenzsystem:

$$\mathbf{R}_{img}^{world} = \begin{pmatrix} \cos(\varphi) \cos(\kappa) & -\cos(\varphi) \sin(\kappa) & \sin(\varphi) \\ \cos(\omega) \sin(\kappa) + \sin(\omega) \sin(\varphi) \cos(\kappa) & \cos(\omega) \cos(\kappa) - \sin(\omega) \sin(\varphi) \sin(\kappa) & \sin(\omega) \cos(\varphi) \\ \sin(\omega) \sin(\kappa) - \cos(\omega) \sin(\varphi) \cos(\kappa) & \sin(\omega) \cos(\kappa) + \cos(\omega) \sin(\varphi) \sin(\kappa) & \cos(\omega) \cos(\varphi) \end{pmatrix}$$

$$\mathbf{R}_{body}^{world} = \mathbf{R}_{img}^{world}$$

Beschreibung der Fehlausrichtung:

$$\mathbf{R}_{sen}^{body} = \begin{pmatrix} \cos(\delta\varphi) \cos(\delta\kappa) & -\cos(\delta\varphi) \sin(\delta\kappa) & \sin(\delta\varphi) \\ \cos(\delta\omega) \sin(\delta\kappa) + \sin(\delta\omega) \sin(\delta\varphi) \cos(\delta\kappa) & \cos(\delta\omega) \cos(\delta\kappa) - \sin(\delta\omega) \sin(\delta\varphi) \sin(\delta\kappa) & \sin(\delta\omega) \cos(\delta\varphi) \\ \sin(\delta\omega) \sin(\delta\kappa) - \cos(\delta\omega) \sin(\delta\varphi) \cos(\delta\kappa) & \sin(\delta\omega) \cos(\delta\kappa) + \cos(\delta\omega) \sin(\delta\varphi) \sin(\delta\kappa) & \cos(\delta\omega) \cos(\delta\varphi) \end{pmatrix}$$

Parametrisierung für Video-Georegistrierung

Drehung vom Navigationskoordinatensystem ins geozentrische kartesische Referenzsystem (WGS 84):

$$\mathbf{R}_{nav}^{WGS84} = \begin{pmatrix} -\sin(\varphi) \cos(\lambda) & -\sin(\lambda) & -\cos(\varphi) \cos(\lambda) \\ -\sin(\varphi) \sin(\lambda) & \cos(\lambda) & -\cos(\varphi) \sin(\lambda) \\ \cos(\varphi) & 0 & -\sin(\varphi) \end{pmatrix}$$

Drehung vom INS-Körpersystem ins Navigationssystem:

$$\mathbf{R}_{body}^{nav} = \begin{pmatrix} \cos(\theta) \cos(\psi) & -\cos(\omega) \sin(\psi) + \sin(\omega) \sin(\theta) \cos(\psi) & \sin(\omega) \sin(\psi) + \cos(\omega) \sin(\theta) \cos(\psi) \\ \cos(\theta) \sin(\psi) & \cos(\omega) \cos(\psi) + \sin(\omega) \sin(\theta) \sin(\psi) & -\sin(\omega) \cos(\psi) + \cos(\omega) \sin(\theta) \sin(\psi) \\ -\sin(\theta) & \sin(\omega) \cos(\theta) & \cos(\omega) \cos(\theta) \end{pmatrix}$$

Drehung vom System der Kameraaufhängung ins INS-Körpersystem:

$$\mathbf{R}_{gimb}^{body} = \begin{pmatrix} \cos(\nu)\cos(\alpha) & -\sin(\alpha) & \sin(\nu)\cos(\alpha) \\ \cos(\nu)\sin(\alpha) & \cos(\alpha) & \sin(\nu)\sin(\alpha) \\ -\sin(\nu) & 0 & \cos(\nu) \end{pmatrix}$$

Drehung vom Bildkoordinatensystem ins System der Kameraaufhängung (beschreibt die Fehlansrichtung):

$$\mathbf{R}_{mount}^{gimb} = \begin{pmatrix} \cos(\Delta\theta)\cos(\Delta\psi) & -\cos(\Delta\omega)s(\Delta\psi) + \sin(\Delta\omega)\sin(\Delta\theta)c(\Delta\psi) & \sin(\Delta\omega)\sin(\Delta\psi) + \cos(\Delta\omega)\sin(\Delta\theta)\cos(\Delta\psi) \\ \cos(\Delta\theta)\sin(\Delta\psi) & \cos(\Delta\omega)\cos(\Delta\psi) + \sin(\Delta\omega)\sin(\Delta\theta)\sin(\Delta\psi) & -\sin(\Delta\omega)\cos(\Delta\psi) + \cos(\Delta\omega)\sin(\Delta\theta)\sin(\Delta\psi) \\ -\sin(\Delta\theta) & \sin(\Delta\omega)\cos(\Delta\theta) & \cos(\Delta\omega)\cos(\Delta\theta) \end{pmatrix}$$

$$\mathbf{R}_{img}^{mount} = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$$

$$\mathbf{R}_{img}^{gimb} = \mathbf{R}_{mount}^{gimb} \cdot \mathbf{R}_{img}^{mount}$$

B Parameterschätzung nach der Methode der kleinsten Quadrate

Ziel einer allgemeinen Messanordnung in der Geodäsie ist es, aus n Beobachtungen \bar{y}_i (Messungen) enthalten im Vektor \mathbf{y} , die m unbekannt Parameter x_j des Vektors \mathbf{x} zu bestimmen, welche über einen funktionalen Zusammenhang mit den Beobachtungen verknüpft sind.

$$\mathbf{y} = F(\mathbf{x}) \quad (\text{B-1})$$

Steht nun eine grössere Anzahl Beobachtungen zur Verfügung, als im Minimum für die Bestimmung der unbekannt Parameter benötigt werden ($m > n$), wird von einem Ausgleichsproblem gesprochen. Für die Lösung einer solchen Aufgabenstellung in der Geodäsie und Photogrammetrie wird vorwiegend das lokale gradientenbasierte Optimierungsverfahren nach Gauss-Newton verwendet (Mautz, 2001). Häufig wird dieses Verfahren in der geodätischen Fachliteratur als Ausgleichung nach Gauss-Markov oder als vermittelnde Ausgleichung bezeichnet. Durch auftretende zufällige Messfehler in den Beobachtungen kann das überbestimmte Gleichungssystem (B-1) nicht diskrepanzfrei gelöst werden, aufgrund dessen das Verbesserungsgleichungssystem

$$\mathbf{y} + \mathbf{v} = F(\mathbf{x}) \quad (\text{B-2})$$

mit dem Verbesserungsvektor \mathbf{v} resultiert. Der eingeführte Vektor \mathbf{v} enthält nun die Verbesserungen zu den einzelnen Beobachtungen, welche eine diskrepanzfreie Lösung gewährleisten. Mittels gradientenbasiertem Optimierungsverfahren lassen sich nun die unbekannt Parameter in dem Sinne bestimmen,

$$\mathbf{v}^T \mathbf{W} \mathbf{v} \Rightarrow \min \quad (\text{B-3})$$

dass die gewichtete Summe der Verbesserungsquadrate zum Minimum wird. Dieses eingeführte Schätzkriterium wird als Methode der kleinsten Quadrate (MdkQ - engl. *Least Squares Method*) oder auch als L2-Norm Optimierung bezeichnet. In der Ausgleichsrechnung werden das funktionale oder auch mathematische Modell, welches den Zusammenhang zwischen Beobachtungen und Unbekannt beschreibt, sowie das stochastische Modell, in welchem die Messgenauigkeiten der einzelnen Beobachtungen modelliert werden, unterschieden. Die Genauigkeiten der zur Verfügung stehenden Beobachtungen werden über die Gewichtsmatrix \mathbf{W} in die Ausgleichung eingeführt und entsprechend der eingeführten Minimumsbedingung (B-3) berücksichtigt. In den

meisten geodätischen und photogrammetrischen Aufgabestellungen ist das mathematische Modell bezüglich der unbekannt Parameter nicht linear und muss für die Schätzung linearisiert werden. Beim Gauss-Newton Verfahren erfolgt die Linearisierung innerhalb der Norm für jede Beobachtung, anhand einer an den Näherungswerten der unbekannt Parameter bis zum Glied 1. Ordnung entwickelten Taylor-Reihe. Woraus nachfolgendes linearisiertes Verbesserungsgleichungssystem folgt.

$$\underbrace{\begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}}_{\mathbf{v}} = \underbrace{\begin{bmatrix} \left(\frac{\partial F_1(\mathbf{x})}{\partial x_1}\right)^0 & \dots & \left(\frac{\partial F_1(\mathbf{x})}{\partial x_n}\right)^0 \\ \vdots & \ddots & \vdots \\ \left(\frac{\partial F_m(\mathbf{x})}{\partial x_1}\right)^0 & \dots & \left(\frac{\partial F_m(\mathbf{x})}{\partial x_n}\right)^0 \end{bmatrix}}_{\mathbf{A}} \cdot \underbrace{\begin{bmatrix} \delta x_1 \\ \vdots \\ \delta x_n \end{bmatrix}}_{\delta \mathbf{x}} - \underbrace{\begin{bmatrix} \bar{y}_1 - y_1 \\ \vdots \\ \bar{y}_m - y_m \end{bmatrix}}_{\mathbf{l}} \quad (\text{B-4})$$

Die verwendete Schreibweise für $(f)^0$ in (B-4) bezeichnet die an der Stelle der Näherungswerte ausgewertete Funktion f . Die Matrix \mathbf{A} wird als Design- oder auch Jacobi-Matrix bezeichnet und enthält die partiellen Ableitungen jeder Beobachtung nach den unbekannt Parametern. Der als Absolutglied benannte Vektor \mathbf{l} umfasst die Differenzen zwischen den Rohmessungen \bar{y}_i und den aus den Näherungswerten der Unbekannt rechnerisch bestimmten Beobachtungen $y_i = (F_i(\mathbf{x}))^0$. In der linearisierten Problemstellung enthält der Vektor $\delta \mathbf{x}$ die gesuchten unbekannt Zuschläge. Im Anschluss ermöglicht eine iterative Lösung nach 'Gauss-Newton' oder 'Levenberg-Marquardt' die Optimierung der Parameter auf Basis der linearisierten Ersatzfunktion ausgehend von extern zur Verfügung gestellten Näherungswerten für die gesuchten unbekannt Parameter. Beide Optimierungsverfahren sind nachfolgend im Detail dargestellt.

B.1 Gauss-Newton Optimierung

Ausgangslage für die Optimierung bildet die Designmatrix \mathbf{A} , der Absolutgliedvektor \mathbf{l} , die Gewichtsmatrix \mathbf{W} sowie die genäherten unbekannt Parameter \mathbf{x}^0 .

Das gesuchte lokale Minimum findet sich durch folgende iterative Lösung:

- i Berechnung der Normalgleichungsmatrix mit anschließender Auflösung nach den unbekannt Zuschlügen:

$$\mathbf{N} = \mathbf{A}^T \mathbf{W} \mathbf{A} \quad (\text{B-5})$$

$$\delta \mathbf{x} = \mathbf{N}^{-1} \mathbf{A}^T \mathbf{W} \mathbf{l} \quad (\text{B-6})$$

- ii Bestimmung der verbesserten unbekannt Parameter:

$$\mathbf{x} = \mathbf{x}^0 + \delta \mathbf{x} \quad (\text{B-7})$$

- iii Test Abbruchkriterium¹

- a) Abbruchkriterium nicht erfüllt: Festlegung verbesserter Näherungswerte mit $\mathbf{x}^0 = \mathbf{x}$ aus welchen \mathbf{A} und \mathbf{l} folgen. Start Berechnung nächste Iteration (\rightarrow i).
- b) Abbruchkriterium erfüllt: Berechnung der resultierenden Genauigkeiten der geschätzten Parameter \mathbf{x} aus

$$\mathbf{v} = \mathbf{A} \cdot \delta \mathbf{x} - \mathbf{l} \quad (\text{B-8})$$

$$m_0 = \sqrt{\frac{\mathbf{v}^T \mathbf{W} \mathbf{v}}{n - m}} \quad (\text{B-9})$$

$$m_{x_j} = m_0 \cdot \sqrt{\mathbf{Q}_{jj}} \quad (\text{B-10})$$

mit $\mathbf{Q} = \mathbf{N}^{-1}$. Darin bezeichnen m_0 und m_{x_j} die Standardabweichungen der Gewichtseinheit und der geschätzten Unbekannt j .

¹ Das Abbruchkriterium kann unterschiedlich definiert werden. Für Details ist auf (Kraus et al., 1996, Kap. 4.4.5) verwiesen.

B.2 Levenberg-Marquardt Optimierung

Zusätzlich zur Ausgangslage bei der Gauss-Newton Optimierung führt dieses Verfahren den Dämpfungsparameter μ , das Verstärkungsverhältnis η sowie die Faktoren τ ¹ und ν ein.

Das Verfahren wird wie folgt initialisiert:

$$\nu = 2$$

$$\mu = \tau \cdot \max(\text{spur}(\mathbf{A}^T \mathbf{W} \mathbf{A})) \quad (\text{B-11})$$

$$F(x) = \mathbf{I}^T \mathbf{W} \mathbf{I} \quad (\text{B-12})$$

Das gesuchte lokale Minimum findet sich durch folgende angepasste iterative Lösung:

- i** Berechnung der erweiterten Normalgleichungsmatrix mit anschließender Auflösung nach den unbekanntem Zuschlägen:

$$\mathbf{N} = (\mathbf{A}^T \mathbf{W} \mathbf{A}) + \mu \mathbf{E} \quad (\text{B-13})$$

$$\mathbf{g} = \mathbf{A}^T \mathbf{W} \mathbf{I} \quad (\text{B-14})$$

$$\delta \mathbf{x} = \mathbf{N}^{-1} \mathbf{g} \quad (\text{B-15})$$

mit der Einheitsmatrix \mathbf{E} .

- ii** Bestimmung der verbesserten unbekanntem Parameter

$$\mathbf{x}_{new} = \mathbf{x}^0 + \delta \mathbf{x} \quad (\text{B-16})$$

aus welchen sich \mathbf{A}_{new} und \mathbf{I}_{new} bilden. Der Dämpfungsparameter bestimmt sich anschliessend wie folgt:

$$F(x_{new}) = \mathbf{I}_{new}^T \mathbf{W} \mathbf{I}_{new} \quad (\text{B-17})$$

$$\eta = \frac{F(x) - F(x_{new})}{\frac{1}{2} \delta \mathbf{x}^T (\mu \cdot \delta \mathbf{x} - \mathbf{g})} \quad (\text{B-18})$$

- a)** Ist $\eta > 0$: setze

$$\mathbf{x}^0 = \mathbf{x}_{new}, \mathbf{A} = \mathbf{A}_{new}, \mathbf{I} = \mathbf{I}_{new} \text{ und } F(x) = F(x_{new})$$

$$\mu = \mu \cdot \max\left\{\frac{1}{3}, 1 - (2\eta - 1)^3\right\} \quad (\text{B-19})$$

$$\nu = 2$$

¹ Für die Anwendung in dieser Arbeit kann für $\tau=10^{-6}$ gewählt werden. (Details vgl. (Madsen et al., 2004, Kap. 3.2))

b) sonst: setze

$$\mu = \mu \cdot \nu \quad (\text{B-20})$$

$$\nu = 2 \cdot \nu \quad (\text{B-21})$$

und beginne mit der Berechnung der nächsten Iteration (\rightarrow **i**)

iii Test Abbruchkriterium¹

c) Abbruchkriterium nicht erfüllt: Start Berechnung nächste Iteration (\rightarrow **i**).

d) Abbruchkriterium erfüllt: Berechnung der resultierenden Genauigkeiten der geschätzten Parameter \mathbf{x}_{new} (vgl. Optimierung nach Gauss-Newton)

¹ Das Abbruchkriterium kann unterschiedlich definiert werden. Für Details ist auf (Madsen et al., 2004, Kap. 3.2) verwiesen.

C Bildorientierung

C.1 Indirekte Georeferenzierung

Die indirekte Georeferenzierung basiert auf der perspektivischen Abbildung, welche die Abbildung eines Objektpunkts \mathbf{P} aus dem übergeordneten Referenzrahmen entlang des Aufnahmestrahls im Bildkoordinatensystem beschreibt (vgl. Abbildung 2-4).

$$\mathbf{b}_{img} = \frac{1}{\lambda} (\mathbf{R}_{img}^{world})^T \cdot (\mathbf{P} - \mathbf{X}_0) \quad (\text{C-1})$$

Die Division der ersten und zweiten Gleichung jeweils durch die Dritte eliminiert den Massstabsfaktor λ , aus welcher die Kollinearitätsgleichungen resultieren:

$$\begin{aligned} x' &= x_0 - c \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \\ y' &= y_0 - c \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \end{aligned} \quad (\text{C-2})$$

Die Skalare r_{ij} in Gl. (C-2) bezeichnen die Elemente der Rotationsmatrix \mathbf{R}_{img}^{world} . Die nachfolgend dargestellten Verfahren 'Einzelbildorientierung' und 'Bündelblockausgleichung' basieren auf diesem funktionalen Zusammenhang.

Einzelbildorientierung (Rückwärtsschnitt)

Ziel dieses Verfahrens ist die Bestimmung der unbekannt Parameter der äusseren Orientierung $X_0, Y_0, Z_0, \omega, \varphi, \kappa$ eines Einzelbildes aus den Bildmessungen \bar{x}^i und \bar{y}^i auf bekannte Passpunkte \mathbf{P}_i . Basierend auf der Kollinearitätsgleichung (C-2) können für die Bildmessungen mit den als bekannt vorausgesetzten Passpunktkoordinaten X_i, Y_i, Z_i und den Parametern der inneren Orientierung c, x_0, y_0 folgende linearisierten Verbesserungsgleichungen aufgestellt werden:

$$\begin{aligned} v_{x'}^i &= dx^{RS} - (\bar{x}^i - (x^i)^0) \\ v_{y'}^i &= dy^{RS} - (\bar{y}^i - (y^i)^0) \end{aligned} \quad (\text{C-3})$$

mit

$$dx^{RS} = \left(\frac{\partial x'}{\partial X_0} \right)^0 dX_0 + \left(\frac{\partial x'}{\partial Y_0} \right)^0 dY_0 + \left(\frac{\partial x'}{\partial Z_0} \right)^0 dZ_0 + \left(\frac{\partial x'}{\partial \omega} \right)^0 d\omega + \left(\frac{\partial x'}{\partial \varphi} \right)^0 d\varphi + \left(\frac{\partial x'}{\partial \kappa} \right)^0 d\kappa$$

$$dy^{RS} = \left(\frac{\partial y'}{\partial X_0} \right)^0 dX_0 + \left(\frac{\partial y'}{\partial Y_0} \right)^0 dY_0 + \left(\frac{\partial y'}{\partial Z_0} \right)^0 dZ_0 + \left(\frac{\partial y'}{\partial \omega} \right)^0 d\omega + \left(\frac{\partial y'}{\partial \varphi} \right)^0 d\varphi + \left(\frac{\partial y'}{\partial \kappa} \right)^0 d\kappa$$

Darin bezeichnen $dX_0, dY_0, dZ_0, d\omega, d\varphi, d\kappa$ die unbekanntes Zuschläge zu den zur Verfügung gestellten Näherungswerten. Die anschliessende Schätzung erfolgt nach der im Anhang B eingeführten vermittelnden Ausgleichung nach der Methode der kleinsten Quadrate.

Bündelblockausgleichung

Im Allgemeinen Fall der Bündelblockausgleichung bilden nun die beiden Bildmessungen \bar{x}'_j und \bar{y}'_j nach Gleichung (C-2) eine Funktion der unbekanntes Parameter der äusseren Orientierung $X_{0j}, Y_{0j}, Z_{0j}, \omega_j, \varphi_j, \kappa_j$ des Bildes j sowie der Verknüpfungspunktkoordinaten X_i, Y_i, Z_i des Objektpunktes i . Daraus resultieren folgende um die partiellen Ableitungen nach den unbekanntes Verknüpfungspunktkoordinaten erweiterten linearisierten Verbesserungsgleichungen:

$$v_{x',j}^i = dx_j^{BA} - (\bar{x}'_j^i - (x'_j)^0)$$

$$v_{y',j}^i = dy_j^{BA} - (\bar{y}'_j^i - (y'_j)^0) \quad (C-4)$$

mit

$$dx_j^{BA} = dx_j^{RS} + \left(\frac{\partial x'}{\partial X} \right)^0 dX_i + \left(\frac{\partial x'}{\partial Y} \right)^0 dY_i + \left(\frac{\partial x'}{\partial Z} \right)^0 dZ_i$$

$$dy_j^{BA} = dy_j^{RS} + \left(\frac{\partial y'}{\partial X} \right)^0 dX_i + \left(\frac{\partial y'}{\partial Y} \right)^0 dY_i + \left(\frac{\partial y'}{\partial Z} \right)^0 dZ_i$$

Dabei bezeichnen $dX_0, dY_0, dZ_0, d\omega, d\varphi, d\kappa$ die unbekanntes Zuschläge der eingeführten Verknüpfungspunkte. Die anschliessende Schätzung erfolgt nach der im Anhang B eingeführten vermittelnden Ausgleichung nach der Methode der kleinsten Quadrate.

C.2 Direkte Georeferenzierung

Die Grundgleichung der direkten Georeferenzierung beschreibt den Zusammenhang zwischen gesuchten Objektpunktkoordinaten und der direkt gemessenen äusseren Orientierung:

$$\mathbf{P} = \mathbf{X}_{world} + \mathbf{R}_{body}^{world} \left(r \cdot \mathbf{R}_{sen}^{body} \cdot \mathbf{e}_{sen} + \Delta \mathbf{S}_{body} - \Delta \mathbf{G}_{body} \right) \quad (\text{C-5})$$

mit Hilfe folgender Terme

$$\mathbf{R}_{body}^{world} = R(\omega, \varphi, \kappa)$$

Rotationsmatrix welche die Drehung vom INS-Körpersystem (Plattformsystem) ins übergeordnete geodätische Referenzsystem beschreibt. (Details zur Parametrisierung der Rotationsmatrix sind in Anhang A zu finden)

$$\mathbf{R}_{sen}^{body} = R(\delta\omega, \delta\varphi, \delta\kappa)$$

Rotationsmatrix welche die Drehung vom Aufnahmesensorkoordinatensystem ins INS-Körperkoordinatensystem beschreibt. (Details zur Parametrisierung der Rotationsmatrix sind in Anhang A zu finden)

$$\mathbf{e}_{sen} = \begin{pmatrix} e_x \\ e_y \\ e_z \end{pmatrix}_{sen}$$

Richtung zum Objektpunkt ausgedrückt im Aufnahmekoordinatensystem als Einheitsvektor.

$$\Delta \mathbf{S}_{body} = \begin{pmatrix} \Delta X_{sen} \\ \Delta Y_{sen} \\ \Delta Z_{sen} \end{pmatrix}_{body}$$

Räumliche Verschiebung des Aufnahmesensorreferenzpunktes bezüglich INS-Referenzpunkt ausgedrückt im Plattformsystem.

$$\Delta \mathbf{G}_{body} = \begin{pmatrix} \Delta X_{GNSS} \\ \Delta Y_{GNSS} \\ \Delta Z_{GNSS} \end{pmatrix}_{body}$$

Räumliche Verschiebung des GNSS-Antennenphasenzentrums bezüglich INS-Referenzpunkt ausgedrückt im Plattformsystem.

r

Distanz zwischen Objektpunkt und Aufnahmesensor-Referenzpunkt.

Wird für die Datenaufnahme ein Flächenbildsensor bzw. eine Videokamera verwendet, sind die beiden Terme wie folgt zu interpretieren:

$$\mathbf{e}_{sen} \Rightarrow \mathbf{b}_{img}$$

Beschreibung des abgebildeten Objektpunktes \mathbf{P} im Bildkoordinatensystem durch die Bildkoordinaten x' und y' .

$$r = c \cdot \lambda \Rightarrow \lambda = \frac{r}{c}$$

Da aus einer Bildaufnahme die Strahllänge r zwischen Projektionszentrum und Objektpunkt unbekannt ist, wird r in Gl. (C-5) durch den Massstabsfaktor λ ersetzt. Der zunächst unbekannte Massstabsfaktor lässt sich durch die Abbildung des Objektpunktes in zwei Bildern bestimmen.

D Zeitdiskretes Kalman-Filter

Ausgangslage

Initiale Annahme des Systemzustandes für die Epoche $t = 0$:

$$\mathbf{x}_0, \mathbf{P}_0 \rightarrow \mathbf{x}_t^{(-)}, \mathbf{P}_t^{(-)}$$

Kenntnis der Kovarianzmatrizen des Systemsrauschens \mathbf{Q}_t und des Messrauschens \mathbf{R}_t .

Schritt 1: Korrektur (update)

Berechnung des Kalman Gain \mathbf{K}_t :

$$\mathbf{K}_t = \mathbf{P}_t^{(-)} \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t^{(-)} \mathbf{H}_t^T + \mathbf{R}_t)^{-1} \quad (\text{D-1})$$

Schätzung des Systemzustandes für die Epoche t aus den zur Verfügung stehenden Beobachtungen \mathbf{z}_t :

$$\mathbf{x}_t^{(+)} = \mathbf{x}_t^{(-)} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \mathbf{x}_t^{(-)}) \quad (\text{D-2})$$

$$\mathbf{P}_t^{(+)} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_t^{(-)} \quad (\text{D-3})$$

Schritt 2: Propagation (prediction)

Propagation des Systemzustandes um Δt auf die nächste Epoche $t+1$:

$$\mathbf{x}_{t+1}^{(-)} = \mathbf{\Phi}_t \mathbf{x}_t^{(+)} \quad (\text{D-4})$$

$$\mathbf{P}_{t+1}^{(-)} = \mathbf{\Phi}_t \mathbf{P}_t^{(+)} \mathbf{\Phi}_t^T + \mathbf{Q}_t \quad (\text{D-5})$$

E Chamfer-Algorithmus

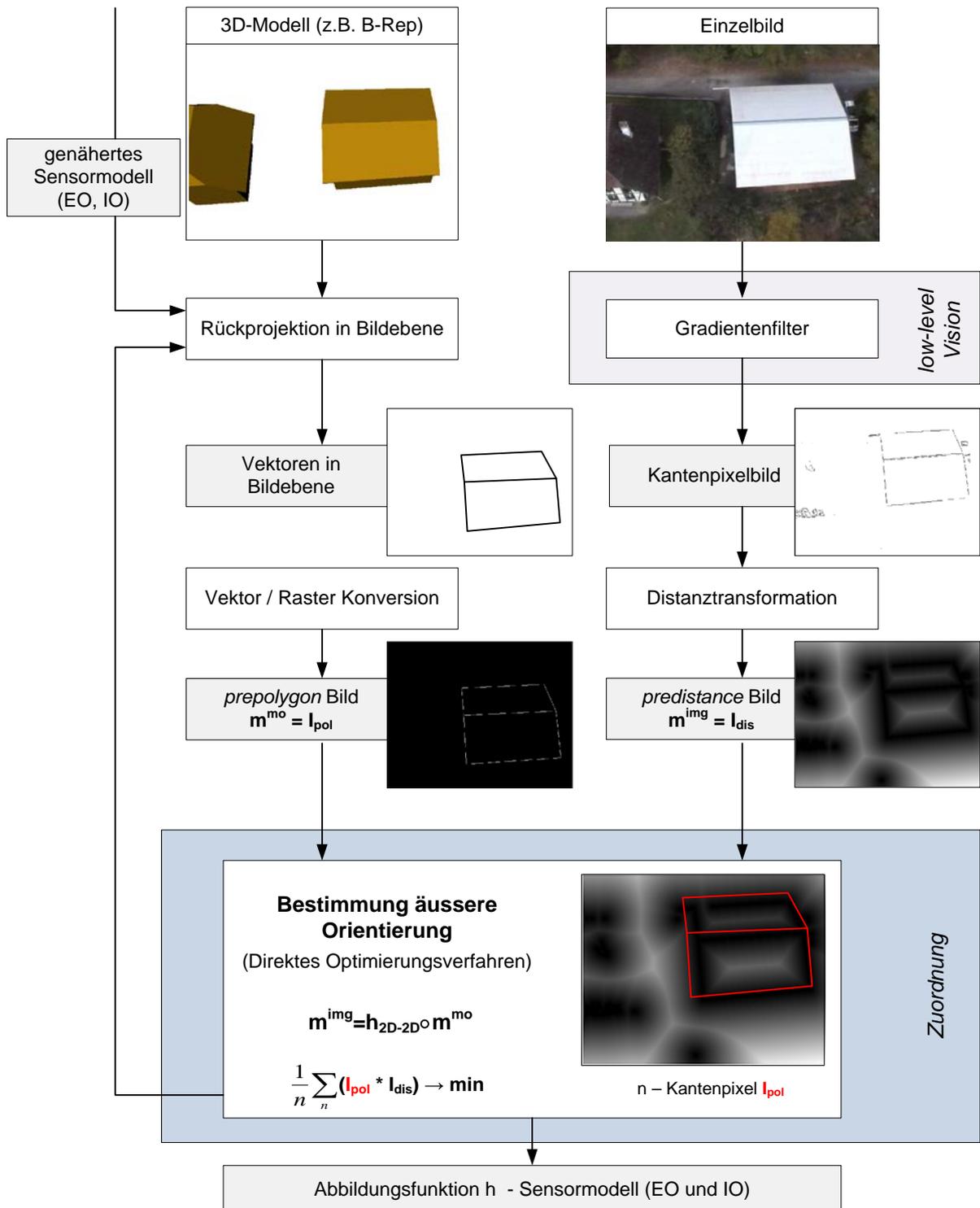


Abbildung E-1: Schematischer Ablauf Chamfer-Algorithmus

Abbildung E-1 zeigt schematisch das Funktionsprinzip des nachfolgend kurz eingeführten Chamfer-Algorithmus. Als erstes werden die Vektoren des Modells mit Hilfe einer näherungsweise vorliegenden äusseren Orientierung der Bildaufnahme in eine virtuelle Bildebene projiziert. Aus den auf diese Weise zurückprojizierten Vektoren entsteht das sogenannte *prepolygon* Bild I_{pol} . Parallel dazu werden aus der Bildaufnahme mittels Gradientenfilter Kantenpixel extrahiert. Ausgehend vom Kantenpixelbild kann beispielsweise mit dem Algorithmus von Borgefors (1986) eine sogenannte Distanztransformation durchgeführt und das *predistance* Bild I_{dis} berechnet werden. Dieses Bild enthält nun in jedem nicht-Kantenpixel die euklidische Distanz zum nächstgelegenen Kantenpixel. Die Zielfunktion bestimmt, aus der elementweisen Multiplikation von I_{pol} mit I_{dis} und anschliessenden Aufsummierung, die Gesamtkosten für die aktuell verwendete äussere Orientierung. Durch schrittweise Veränderung der einzelnen äusseren Orientierungsparameter werden mit Hilfe des hier beschriebenen direkten Optimierungsverfahren die Gesamtkosten minimiert und damit die bestmögliche Übereinstimmung gefunden.

F A* Algorithmus - informierte Baumsuche

Pseudoquelltextdarstellung der A* Suche für Maximierung der Evaluationsfunktion:

```

Algorithmus aStar( $m^{mo}$ ,  $m^{img}$ )

//Gesuchte Zuordnungsliste
matchList := null;

//Anlegen des Wurzelknotens
rootNode := createNode(0,0,0,0,0, create(hStack));

//Initialisierung des Knotenregisters
create(nodeStack);
nodeStack.add(rootNode);

iterate := true;

while iterate do
    //Extrahiere den Knoten mit dem höchsten geschätzten Nutzen
    exNode := nodeStack.getFirst();

    nodes := expandNode(exNode,  $m^{mo}$ ,  $m^{img}$ );

    if nodes != null then
        nodeStack.add(nodes);

        //Sortieren des Registers absteigend nach der
        //Variable  $E^*(n)$ 
        sort(nodeStack);
    else
        iterate := false;
        matchList := exNode.hStack;
    fi;
od;

```

```

function [nodeList] = expandNode(Node, modDesA, modDesB)

h := Node.hStack;
[pA, pB] := h.getLast();

if pA+1 ∈ modDesA then
    create(nodeList);

    //Bestimmung der möglichen Zuordnungen für Primitive pA+1
    list_pB := createPossibleMatches(pA+1, h);

    while list_pB.hasMore() do
        pB := list_pB.next();

        //Berechnung des Nutzens aus den Modellbeschreibungen und
        //der bisherigen Zuordnung
        g_n := estimate_g(n)(pA+1, pB, modDesA, modDesA, h);
        hs_n := estimate_h*(n)(pA+1, pB, modDesA, modDesA, h);

        nodeList.add(createNode(pA+1, pB, g_n, hs_n, h));
    od;
else
    nodeList := null;
fi;

function [Node] = createNode(pA, pB, g_n, hs_n, h)

//Zuordnungsliste
Node.hStack := h;
Node.hStack.add(match(pA, pB));
//Nutzen bis zum Knoten n
Node.g(n) := g_n;
//Schätzung des weiter anfallenden Nutzens
Node.h*(n) := hs_n;
//Geschätzter Nutzen der Evaluationsfunktion
Node.E*(n) := g_n + hs_n;

```

G Kontinuierlicher Bündelblockausgleich

Pseudoquelltextdarstellung der kontinuierlichen Bündelblockausgleichung:

Algorithmus kontinuierlicherBündelblock(n)

$a := 0;$

while liegt neues \mathbf{SB}_{RGB} an **do**

$a := a+1;$

 // Aufbau Datenstruktur für Schlüsselbild a

 create($\mathbf{eo}_a^{\text{app}} := \mathbf{EO}_{\text{app}}$);

 //Überführt alle VP-Beobachtungen des Schlüsselbildes a
 //aus Merkmalsverfolgung

$\mathbf{B}_a^{j \in \text{SB}_a} := \text{getTrackedVP}(\mathbf{SB}_{\text{RGB}});$

for $z = 1$ **to** length($\mathbf{B}_a^{j \in \text{SB}_a}$) **do**

$j := \mathbf{B}_a[z];$

if $\mathbf{vp}_j \in \mathbf{VP}$ **then**

${}^{\text{obs}+1}\mathbf{vp}_j;$

else

 create(${}^1\mathbf{vp}_j^{\text{nip}}$);

fi;

od;

 //Anwendung Bild-zu-3D-Stadtmodellzuordnung für Schlüsselbild a

$\mathbf{P}_a^{k \in \text{SB}_a}, \mathbf{E}_a^{l \in \text{SB}_a}, \mathbf{RP}_a^{k \in \text{SB}_a}, \mathbf{RE}_a^{j \in \text{SB}_a} := \text{imgToModMatch}(\mathbf{SB}_{\text{RGB}}, \mathbf{EO}_{\text{app}}, \mathbf{IO}_{\text{DG}});$

 // Löschen Datenstruktur für Schlüsselbild $a-n$

if $a > n$ **then**

for $z = 1$ **to** length(\mathbf{B}_{a-n}) **do**

$j := \mathbf{B}_{a-n}[z];$

if ${}^{>1}\mathbf{vp}_j$ **then**

${}^{\text{obs}-1}\mathbf{vp}_j;$

else

 delete(\mathbf{vp}_j);

fi;

od;

```

        delete(  $\mathbf{e}o_{a-n}$  ,  $\mathbf{B}_{a-n}$  ,  $\mathbf{P}_{a-n}$  ,  $\mathbf{E}_{a-n}$  ,  $\mathbf{RP}_{a-n}$  ,  $\mathbf{RE}_{a-n}$  );
    fi;

    // Blockausgleichung für Bündel a
    baB := B ;
    for z =1 to length( VP ) do
        vpj := VP [z];
        if vpj ∉  $\geq^2 \mathbf{VP}^{\neq nip}$  then
            j := vpj;
            delete( baBj )
        fi;
    od;
    EOadj , VPadj := BAdjustment( EO , baB , bRP , eRE ,  $\geq^2 \mathbf{VP}^{\neq nip}$  , RP , RE );

    // Initialisierung VP
    for z =1 to length(  $\geq^2 \mathbf{VP}^{nip}$  ) do
        j :=  $\geq^2 \mathbf{VP}^{nip}$  [z];
        obsvpjapp := intersectVP( EO , Bj );
    od;
od;

```

Dank

Für die ausgezeichneten Arbeitsbedingungen und die grosse Unterstützung zum Gelingen dieser Arbeit möchte ich mich bei der Fachhochschule Nordwestschweiz und insbesondere bei dem Institut Vermessung und Geoinformation herzlich bedanken. Ein grosser Dank geht an Prof. Dr. Ralf Reulke und an Prof. Dr. Stephan Nebiker für die Betreuung dieser Arbeit und die vielen wertvollen Hinweise, Tipps und Diskussionen. Die Zusammenarbeit hat mir immer viel Spass bereitet und ich konnte dabei sehr viel lernen und profitieren.

Herzlich bedanke ich mich bei: Adrian Annen, Kevin Flückiger, Susanne Bleisch, Martin Christen, Andreas Barmettler, Joel Burkhard und Reto Stibler, für die Unterstützung bei der Durchführung der unzähligen Drohnenflüge sowie für die vielen fachlichen Inputs, spannenden Diskussionen und die ausgezeichnete Zusammenarbeit.

Speziell bedanke ich mich bei meinen Freunden und bei meiner Familie die mich bei dieser Arbeit immer unterstützt und mir auch manchmal die dafür zusätzliche Motivation gegeben haben.

“Geodätisch-geophysikalische Arbeiten in der Schweiz”
(Fortsetzung der Publikationsreihe ”Astronomisch-geodätische Arbeiten in der Schweiz”)
der Schweizerischen Geodätischen Kommission (www.sgc.ethz.ch)

- 63 2001 Analysis of refraction influences in geodesy using image processing and turbulence models. Philipp Flach, 175 Seiten.
- 64 2003 INS/GPS Integration for Pedestrian Navigation. V. Gabaglio, 161 Seiten.
- 65 2003 Efficient Methods for Determining Precise Orbits of Low Earth Orbiters Using the Global Positioning System. Heike Bock, 214 Seiten.
- 66 2003 Capteurs et Algorithmes pour la Localisation Autonome en Mode Pédestre. Quentin Ladetto, 121 Seiten.
- 67 2004 GPS based Determination of the Integrated and Spatially Distributed Water Vapor in the Troposphere. Marc Troller, 172 Seiten.
- 68 2005 Geodetic Mobile Solar Spectrometer. Alexander Somieski, 205 Seiten.
- 69 2005 Absolute Airborne Gravimetry. Henri Baumann, 142 Seiten.
- 70 2006 The *Swiss Trolley* – A Modular System for Track Surveying. Ralph Glaus, 184 Seiten.
- 71 2006 Development of a Robotic Mobile Mapping System by Vision-Aided Inertial Navigation: A Geomatics Approach. Fadi Atef Bayoud, 157 Seiten.
- 72 2007 Das neue Landeshöhenetz der Schweiz LHN95. Andreas Schlatter, 373 Seiten.
- 73 2007 Pseudo-Stochastic Orbit Modeling of Low Earth Satellites Using the Global Positioning System. Adrian Jäggi, 214 Seiten.
- 74 2008 Cartographie Mobile en Temps Réel. Hervé Gontran, 162 Seiten.
- 75 2008 Mutual Validation of Satellite-Volume 75 Geodetic Techniques and its Impact on GNSS Orbit Modeling. Claudia Flohrer, 198 Seiten.
- 76 2009 High-resolution GPS tomography in view of hydrological hazard assessment. Simon Lutz, 202 Seiten.
- 77 2009 Trajectory Determination and Analysis in Sports by Satellite and Inertial Navigation. Adrian Wägli, 175 Seiten.
- 78 2009 GPS Based Dynamic Monitoring of Air Pollutants in the City of Zurich. Philippe Kehl, 156 Seiten.
- 79 2010 In-flight Quality Assessment and Data Processing for Airborne Laser Scanning. Philipp Schaer, 166 Seiten.
- 80 2010 Sea Surface Topography and Marine Geoid by Airborne Laser Altimetry and Shipborne Ultrasound Altimetry. Philippe Limpach, 207 Seiten.
- 81 2010 Global Gravity Field Determination Using the GPS Measurements Made Onboard the Low Earth Orbiting Satellite CHAMP. Lars Prange, 213 Seiten.
- 82 2011 Analysis of long-term GPS observations in Greece (1993-2009) and geodynamic implications for the Eastern Mediterranean. Michael D. Müller, 186 Seiten.
- 83 2011 Combined Analysis of Observations from Different Global Navigation Satellite Systems. Michael Meindl, 150 Seiten.
- 84 2012 Water Vapor Tomography using Global Navigation Satellite Systems. Donat Perler, 198 Seiten. 2012
- 85 2012 Echtzeit-Georegistrierung von Videodaten mit Hilfe von Navigationssensoren geringer Qualität und digitalen 3D-Landschaftsmodellen, Hannes Eugster, 182 Seiten.