Die approbierte Originalversion dieser Diplom-/Masterarbeit ist in der Hauptbibliothek der Technischen Universität Wien aufgestellt und zugänglich.
http://www.ub.tuwien.ac.at

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology.
http://www.ub.tuwien.ac.at/eng

TECHNISCHE UNIVERSITÄT WIEN

D I P L O M A R B E I T

Traufenmodellierung mittels Matching von Liniensegmenten

Ausgeführt am

Department für Geodäsie und Geoinformation
Forschungsgruppe Photogrammetrie

der Technischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Josef Jansa
Univ.Lektor Dipl.-Ing. Dr.techn. Robert F. Tobler

durch

Gernot Felfernig, BSc
1120 Wien, Tivoligasse 43/2

Wien, Mai 2015 Gernot Felfernig

Kurzfassung

Zur Erstellung genauer Orthophotos werden exakte Digitale Oberflächenmodelle benötigt. Vor allem an Traufenkanten treten aber häufig Ungenauigkeiten auf, welche meist auf Abrundungseffekte, die bei der Interpolation zur Berechnung von Digitalen Oberflächenmodellen auftreten, zurückzuführen sind und sich im Orthophoto durch Reliefverschiebungen und unscharfe Abbildung der Traufen zeigen. Daher wurden zwei Methoden hinsichtlich ihrer Eignung, Traufenkanten genau und automatisch zu modellieren und damit Digitale Höhenmodelle zu verbessern, untersucht, wobei eine davon im Rahmen dieser Masterarbeit eigenständig entwickelt wurde. Die Methoden verwenden orientierte Luftbilder und basieren auf demselben Prinzip, dabei beschränken sie sich nicht nur auf Traufen, sondern sollen alle sichtbaren Kanten rekonstruieren. Zunächst werden die abgebildeten Kanten in den einzelnen Luftbildern als Liniensegmente detektiert. Dann sollen im Zuge eines Matchings der Liniensegmente korrespondierende Kanten in den verschiedenen Luftbildern gefunden werden, um diese anschließend durch einen Vorwärtsschnitt dreidimensional zu rekonstruieren, wobei bei drei oder mehr korrespondierenden Liniensegmenten ein überbestimmter Vorwärtsschnitt vorliegt, der einer Ausgleichung unterzogen werden muss. Die Untersuchungen zeigen, dass beide Methoden für die getesteten Luftbilder, welche eine typische Szene im urbanen Raum abbilden, keine verlässlichen Ergebnisse liefern. Abschließend wird anhand von Beispielen erläutert, woran die beiden Algorithmen scheitern und damit aufgezeigt, wo sie an ihre Grenzen stoßen.

Abstract

For the generation of accurate orthophotos exact digital surface models are needed. A common problem is the rounding effect that occurs during the interpolation of digital surface models and leads to inaccurate eave lines and relief displacements in orthophotos. Therefore, two methods for the automatic and exact reconstruction of eaves and thus for the improvement of digital elevation models were determined in the context of this master thesis. One of them was specially developed for this research. Both methods are not restricted to eaves, but aim to reconstruct all visible edges. They need oriented aerial images as input and operate on the same principle: First edges are detected as line segments in all aerial images. Then a line segment matching is performed in order to find correspondences, which are reconstructed as three-dimensional edges by spatial intersection afterwards. If there are three or more corresponding line segments, an adjustment of differences has to be performed. Investigations of this master thesis reveal that both methods do not yield reliable results for the given aerial images that are shot in a typical urban scene. Finally, the reasons for their failure are shown by examples and thus the limitations of the algorithms are outlined.

Danksagung

Auf dem Weg zur Fertigstellung dieser Masterarbeit bin ich von mehreren Menschen begleitet worden, denen ich an dieser Stelle herzlich dafür danken möchte.

An erster Stelle danke ich meinen Eltern, Anna und Johann Felfernig, und meiner Schwester MMag. Marlies Felfernig, für ihre Unterstützung und Wertschätzung während meines Studiums.

Des Weiteren gebührt mein Dank dem Zentrum für Virtual Reality und Visualisierung, welches mir ermöglicht hat, meine Masterarbeit im Rahmen eines interessanten, praxisbezogenen Projekts durchzuführen. Vor allem mein Co-Betreuer Univ.Lektor Dipl.-Ing. Dr.techn. Robert F. Tobler sowie Dipl.-Ing. Dr.techn. Stefan Maierhofer haben mich dabei unterstützt und durch ihre Anregungen und Impulse die Richtung meiner Masterarbeit entscheidend mitbeeinflusst.

Nicht zuletzt möchte ich mich bei jenen Wissenschaftlern am Department für Geodäsie und Geoinformation der TU Wien bedanken, die mich während meines interessanten und abwechslungsreichen Studiums geprägt haben. Im Zusammenhang mit meiner Masterarbeit möchte ich dabei Senior Scientist Dipl.-Ing. Dr.techn. Camillo Ressl hervorheben, der mir bei der Lösung des Ausgleichungsproblems geholfen hat und natürlich meinen Betreuer Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Josef Jansa, der immer ein offenes Ohr für meine Fragen hatte und mir mit Rat zur Seite stand.

Inhaltsverzeichnis

1 Einleitung 1

2 Aufgabenstellung 2

2.1 Aufbau der Arbeit 2

3 Grundlagen 3

3.1 Luftbild 3

3.2 Digitales Höhenmodell (DHM) 5

3.2.1 Digitales Geländemodell (DGM) 5

3.2.2 Digitales Oberflächenmodell (DOM) 5

3.3 Orthophoto (OP) 6

3.3.1 Genauigkeit des Orthophotos 7

3.3.1.1 Reliefverschiebung 7

3.4 Matching von Liniensegmenten 9

3.4.1 Vorgangsweise 9

3.4.2 Herausforderungen 10

3.4.3 Einteilung von Ansätzen 10

4 Stand der Forschung 13

4.1 Matching von Liniensegmenten 13

4.1.1 Matching von Liniensegmenten nach Bay et al. (2005) 13

4.1.1.1 Farbhistogramm-basiertes Matching 13

4.1.1.2 Topologischer Filter 15

4.1.1.3 Bestimmung weiterer Korrespondenzen 16

4.1.2 Matching von Liniensegmenten nach Schmid und Zisserman (1997) ... 17

4.1.2.1 Deskriptor 17

4.1.2.2 Ansatz für kurze Basis 17

4.1.2.3 Ansatz für lange Basis 18

4.1.2.4 Erweiterung auf drei Bilder 19

4.1.3 Implementierung zum Matching von Liniensegmenten und zur Rekonstruktion von 3D-Kanten von Werner und Zisserman (2002) 21

4.1.3.1 Algorithmus 21

5 Daten 23

5.1 Luftbilder 23

5.2 Digitales Oberflächenmodell 23

6 Methoden 24

6.1 Detektion von Liniensegmenten 25

6.1.1 Clustering der Liniensegmente 25

6.2 Matching von Liniensegmenten 26

6.2.1 Spektraler Merkmalsraum 26

6.2.2 Nebenbedingungen 28

6.2.2.1 Nebenbedingung 1 − innerhalb des Suchfensters 28

6.2.2.2 Nebenbedingung 2 − kein schleifender Schnitt 32

6.2.2.3 Abschließende Bemerkungen zum Matching der Liniensegmente 33

6.3 Rekonstruktion von 3D-Kanten 35

6.3.1 Ausgleichung des Vorwärtsschnittes 35

6.3.2 Erkennen und Eliminieren von Ausreißern 38

6.3.3 Bestimmung von 3D-Kanten 41

7 Ergebnisse 42

7.1 Ergebnisse mittels der entwickelten Implementierung 42

7.1.1 Symbolik der Abbildungen 42

7.1.2 Beispiel 1 – ähnliche sich wiederholende Strukturen 42

7.1.3 Beispiel 2 – Richtung der Liniensegmente 43

7.1.4 Beispiel 3 – Farbprofil nur auf einer Seite vergleichbar 44

7.1.5 Beispiel 4 – Doppelkanten 45

7.1.6 Beipsiel 5 – Probleme bei der Kantendetektion 47

7.1.7 Beispiel 6 – erfolgreiches Matching und 3D-Rekonstruktion 48

7.2 Ergebnisse mittels der Implementierung von Werner und Zisserman (2002) 49

8 Zusammenfassung und Ausblick 52

9 Literaturverzeichnis 53

1 Einleitung

Die Entwicklung in der Computertechnologie und damit auch in der graphischen Datenverarbeitung sowie die zunehmende Verfügbarkeit von qualitativ hochwertigen digitalen Daten haben zu neuen Formen der Aufgabenbearbeitung in der öffentlichen Verwaltung und in der privaten Wirtschaft geführt. Damit einher geht auch die wachsende Nachfrage nach digitalen topographischen Informationen, die als Basisdaten für digitale raumbezogene thematische Bearbeitungen unverzichtbar geworden sind. Um den steigenden Bedarf an genauen, einfach und schnell zu erfassenden, günstigen und großmaßstäbigen Geo-Informationen unserer Umwelt von verschiedenen Anwendern Geographischer Informationssysteme (GIS) wie z. B. die öffentlichen Stellen, Ingenieurbüros, Versorgungsunternehmen und viele mehr für Planungsaufgaben, Resourcenverwaltung und Umweltmonitoring etc. befriedigen zu können, müssen in immer kürzeren Zeitabständen aktuelle Grundlagendaten in digitaler Form zu fairen und verbraucherfreundlichen Preisen zur Verfügung gestellt werden. Als solche Grundlagendaten bieten sich Digitale Orthophotos an, die durch eine entsprechende Transformation aus Luftbildern oder Satellitenbildern mit Zentralprojektion zu einer georeferenzierten orthogonalen Photokarte umgeformt werden. Durch die Georeferenzierung wird ein Raumbezug gewährleistet und eine spätere gemeinsame Nutzung verschiedener Datensätze im GIS vereinfacht.

Digitale Orthophotos weisen viele Vorteile auf: (i) sie liefern die hohe Genauigkeit und den großen Informationsgehalt der Luftbilder; (ii) sie sind aufgrund der Möglichkeit zur Erfassung großer Flächen in kurzen Zeitabständen als homogene, flächendeckende Datensätze verfügbar; (iii) die kurze Produktionszeit und die niedrigen Herstellungskosten führen zu einem sehr guten Preis- Leistungsverhältnis; (iiii) sie bieten Möglichkeiten zur (automatischen) computergestützten Informationsextraktion und Ableitung von Folgeprodukten; (iiiii) und lassen ihren Nutzen durch Integration von Zusatzinformation und Kombination mit bestehenden Datensätzen steigern [Kersten, 1996].

Daher dienen Digitale Orthophotos in zunehmendem Maße als wichtige und aktuelle Grundlage für viele Anwendungen in Geographischen Informationssystemen. Durch die Verfügbarkeit von immer höher auflösenden digitalen topographischen Datensätzen steigen allerdings gleichzeitig auch die Anforderungen an die Qualität und Genauigkeit Digitaler Orthophotos. Dabei kommen auch gewisse Schwächen und Anwendungsgrenzen zum Vorschein. Da das Digitale Orthophoto ein Folgeprodukt aus Luftbild und Digitalem Oberflächenmodell (DOM) ist, hängt seine Genauigkeit von jener der beiden ihm zugrunde liegenden Datensätze ab. Ein bekanntes, häufig auftretendes Problem in Orthophotos resultiert aus Ungenauigkeiten von Traufenkanten im zugrunde liegenden Digitalen Oberflächenmodell und äußert sich durch Reliefverschiebungen und ungenaue bzw. falsche Abbildung dieser Traufenkanten.

Daher ist es das Ziel dieser Masterarbeit, Methoden und Ansätze zur automatischen Modellierung von Traufenkanten zu untersuchen und zwar hinsichtlich ihrer Eignung, Digitale Oberflächenmodelle oder allgemeiner − Digitale Höhenmodelle (DHM) − zu verbessern und damit genauere Digitale Orthophotos zu generieren.

2 Aufgabenstellung

Ziel dieser Masterarbeit ist es einen Ansatz zur Erstellung möglichst genauer Orthophotos zu untersuchen und zu testen. Durchgeführt werden diese Untersuchungen im Rahmen eines Projekts am Zentrum für Virtual Reality und Visualisierung (VRVis) in Zusammenarbeit mit der Firma Meixner Imaging GmbH. Für dieses Projekt stehen orientierte Luftbilder und ein Digitales Oberflächenmodell von Graz zur Verfügung, mit dem Ziel, möglichst genaue Orthophotos mit dem Hauptaugenmerk auf der Beseitigung von Ungenauigkeiten an Traufenkanten zu produzieren. Ein häufig auftretendes Problem, welches auch bei dem verwendeten Digitalen Oberflächenmodell vorkommt, sind Ungenauigkeiten an den Traufenkanten, welche nicht scharf abgebildet, sondern aufgrund der notwendigen Interpolation bei der Berechnung des Digitalen Oberflächenmodells abgerundet dargestellt sind. Diese Abrundung der Traufenkanten führt bei der Berechnung eines Orthophotos zu Fehlern, die sich durch Reliefverschiebungen und unscharfe Abbildung der Traufenkanten äußern.

Daher wird im Rahmen dieser Masterarbeit eine Methode zur genauen und automatischen Modellierung von Traufenkanten untersucht. Bei dieser Methode sollen zunächst die abgebildeten Traufenkanten in den einzelnen Luftbildern als Liniensegmente detektiert werden. Aufgrund der hohen Überlappung der Luftbilder sind die Traufenkanten meistens in mehreren Luftbildern sichtbar. Im nächsten Schritt sollen im Zuge eines Matchings der Liniensegmente korrespondierende Traufenkanten in den verschiedenen Luftbildern gefunden werden. Danach sollen die Traufenkanten durch Vorwärtsschnitt der korrespondierenden Liniensegmente dreidimensional rekonstruiert werden, wobei bei drei oder mehr korrespondierenden Liniensegmenten ein überbestimmter Vorwärtsschnitt vorliegt, der einer Ausgleichung unterzogen werden muss. Zuletzt soll festgestellt werden, inwiefern diese Methode der Modellierung der Traufenkanten zur Verbesserung von Digitalen Oberflächenmodellen und somit auch zur Berechnung genauerer Digitaler Orthophotos genutzt werden kann.

2.1 Aufbau der Arbeit

Im anschließenden Kapitel 3 werden grundlegende Informationen über Luftbilder und deren Orientierung, über Orthophotos und inwiefern sich Fehler im zugrunde liegenden Höhenmodell auswirken, sowie über das Matching von Liniensegmenten dargelegt. In Kapitel 4 werden bestehende Ansätze und eine Implementierung zum Matching von Liniensegmenten und zur anschließenden dreidimensionalen Rekonstruktion dieser Liniensegmente vorgestellt, welche dem aktuellen Stand der Forschung entsprechen. Die verwendeten Daten werden in Kapitel 5 beschrieben und in Kapitel 6 wird die entwickelte Methodik vorgestellt, welche teilweise auf den Ansätzen aus Kapitel 4 beruht. In Kapitel 7 werden die Ergebnisse gezeigt, verglichen und interpretiert, welche einerseits mit der Methodik aus Kapitel 5 und andererseits mit der Implementierung aus Kapitel 4 erzielt wurden. Abschließend wird eine Zusammenfassung in Kapitel 8 und ein Ausblick, der Ideen zur Verbesserung der Ansätze enthält, in Kapitel 9 gegeben.

3 Grundlagen

3.1 Luftbild

Ein wesentlicher Teil der aus flugzeuggestützter Fernerkundung gewonnenen Daten wird mit Photokameras aufgezeichnet und kann als Bilder (Luftbilder) dargestellt werden.

Diese enthalten Informationen über das Reflexionsverhalten der aufgenommenen Objekte in dem vom Sensor aufgezeichneten Bereich des elektromagnetischen Spektrums (Kanal). Diese Information wird bei digitalen Bildern in einer Matrix (Bildmatrix) gespeichert. In dieser Bildmatrix wird die Position jedes Bildpunkts, auch Pixel genannt, eindeutig durch seine Bildkoordinaten festgelegt. Ein Multispektralbild besteht aus mehreren Einzelbildern, die in unterschiedlichen Kanälen aufgenommen wurden. Bei einem Multispektralbild, das aus *n* Einzelbildern besteht, hat also jedes Pixel *n* Werte (Grauwerte) [Kulbach, 1997].

Geometrisch wird diese Abbildung in erster Annäherung durch eine Zentralprojektion beschrieben. Die Zentralprojektion wird durch die Bildebene Π und das Projektionszentrum *O*, das nicht in Π liegen darf, definiert. Der Schnittpunkt des Projektionsstrahles *OP* mit Π ist das Zentralperspektive-Bild *P'* des Punktes *P*. Die Gleichungen der Zentralprojektion beschreiben die Transformation des Objektpunktes *P* in den Bildpunkt *P'*. Die Abbildung 3.1 zeigt den Zusammenhang in einer Skizze.

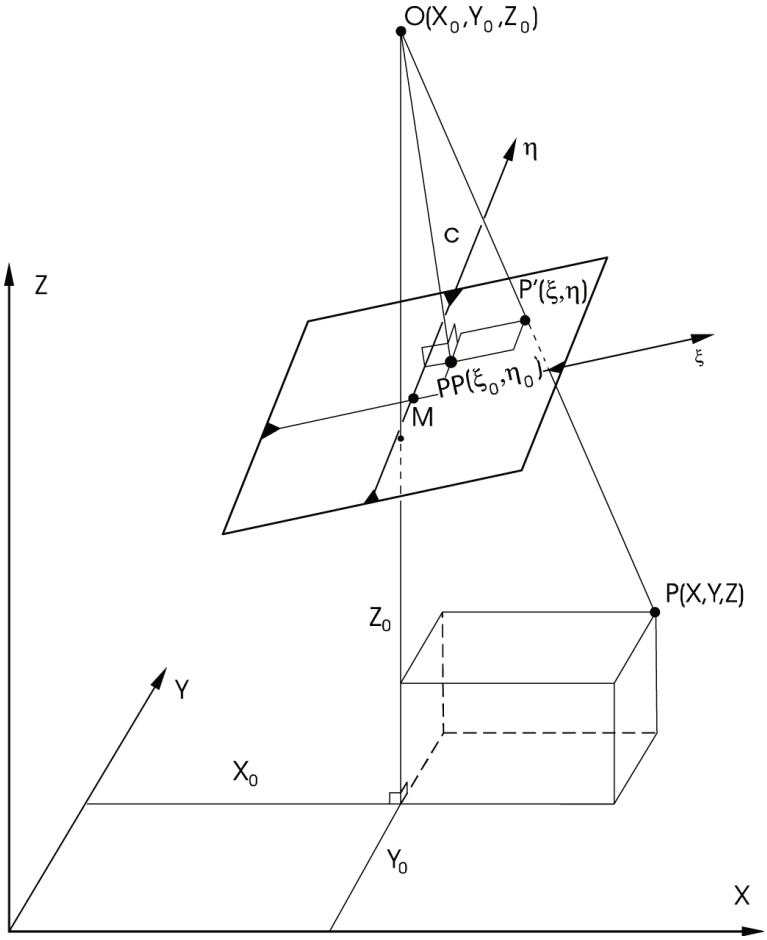

Abbildung 3.1: Zusammenhang zwischen Bild- und Objektkoordinaten bei einer Zentralprojektion [Kraus, 2004].

Zur mathematischen Formulierung dieses Vorganges legen wir zwei räumliche kartesische Koordinatensysteme fest: Das Objektkoordinatensystem *X-Y-Z,* das zur Beschreibung des räumlichen Objektes (Gelände) dient und das Bildkoordinatensystem ξ-η-*z,* dessen $\xi\eta$-Ebene in die Bildebene Π fällt.

Die Koordinaten des Projektionszentrums in der $\xi\eta$-Ebene definieren den Hauptpunkt PP (ξ_0, η_0, 0). Der Abstand des Hauptpunktes vom Projektionszentrum wird als Kammerkonstante c bezeichnet. Die Bildkoordinaten des Projektionszentrums O (ξ_0, η_0, c) sind die Parameter der inneren Orientierung des Bildes. Die äußere Orientierung ist durch die Objektkoordinaten des Projektionszentrums O (X_0, Y_0, Z_0) und durch die Stellung des Bildsystems gegenüber dem Objektsystem festgelegt. Diese Stellung wird durch eine Rotationsmatrix R beschrieben. Innere und äußere Orientierung definieren die Parameter der Zentralprojektion. Die Abbildungs- oder Kollinearitätsgleichungen eines Objektpunktes *P (X, Y, Z)* lauten [Kraus, 2004]:

$$\xi = \xi_0 - c\frac{r_{11}(X-X_0)+r_{21}(Y-Y_0)+r_{31}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)}$$
$$\eta = \eta_0 - c\frac{r_{12}(X-X_0)+r_{22}(Y-Y_0)+r_{32}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)} \qquad (3.1)$$

mit	ξ, η	$\cdots$	Bildkoordinaten des verzerrten Luftbildes
	ξ_0, η_0	$\cdots$	Bildkoordinaten des Bildhauptpunktes PP
	X, Y, Z	$\cdots$	Koordinaten des Objektpunktes
	X_0, Y_0, Z_0	$\cdots$	Koordinaten des Aufnahmezentrums
	r_{ik}	$\cdots$	Elemente der Rotationsmatrix
	c	$\cdots$	Kamerakonstante

3.2 Digitales Höhenmodell (DHM)

In der Fernerkundung und Photogrammetrie verwendet man häufig DHMe als vereinfachende bildliche oder mathematische Darstellung zur Beschreibung der Erdoberfläche. Diese Modelle sind nicht Selbstzweck, sondern ein Zwischenschritt für bestimmte Anwendungen (Höhenschichtlinien für topographische Karten, Überschwemmungssimulationen usw.) [Pfeifer, 2003]. Die Höheninformation, die das DHM darstellt, kann sich unter anderem auf das Gelände, dann spricht man von einem Digitalem Geländemodell (DGM), oder auf die Oberfläche, in diesem Fall nennt man es Digitales Oberflächenmodell (DOM), beziehen. Formal lässt sich ein Höhenmodell typischerweise durch eine Funktion $Z = f(X, Y)$ definieren. Dabei wird jedem Bodenpunkt (definiert durch X und Y) eine Höhe Z zugeordnet. Weil dadurch keine echte Dreidimensionalität unterstützt wird, spricht man auch von $2\frac{1}{2}$-D-Repräsentation. Rastermodelle zum Beispiel böten auch die Möglichkeit einer echten 3-D-Repräsentation, aber für Höhenmodelle ist der zuvor erwähnte Ansatz in den meisten Fällen ausreichend. Die Verspeicherung und Repräsentation des Modells erfolgt digital, wobei es verschiedene gebräuchliche Datenstrukturen gibt (Dreiecksvermaschung, Gittermodell, Rastermodell, Hybrides Geländemodell) (siehe [Kraus, 2000]). Die am häufigsten verwendeten DHMe werden nachfolgend kurz beschrieben.

3.2.1 Digitales Geländemodell (DGM)

Das DGM beschreibt die Erdoberfläche unter Vernachlässigung aller künstlichen Objekte (z. B. Gebäude, Brücken) und des natürlichen Bewuchses (z. B. Bäume, Sträucher). Als Höhe wird hier also die Geländehöhe genommen, wobei man unter Gelände die Grenzschicht vom gewachsenen Boden zur Luft versteht.

3.2.2 Digitales Oberflächenmodell (DOM)

Das DOM beschreibt die Erdoberfläche ähnlich dem DGM, aber inklusive der Beschreibung von künstlichen und natürlichen Objekten (Gebäude, Baumkronen usw.). Es wird also die Höhe der, aus der Vogelperspektive gesehen, obersten sichtbaren Fläche genommen.

3.3 Orthophoto (OP)

Die Original-Luftbilder enthalten zentral-perspektive Verzerrungen. Um die Textur (= Farbinformation) des Luftbildes mit anderen Datenbeständen in Beziehung setzen bzw. vergleichen zu können, ist es sinnvoll, eine Entzerrung des Bildinhalts durchzuführen, sodass die Geometrie überall einer Orthogonalprojektion entspricht. Das Ergebnis ist ein sogenanntes Orthophoto, welches geometrisch korrekt ist, also keine Verzerrungen beinhaltet, und einen einheitlichen Maßstab hat. Zur OP-Berechnung wird (um die Aufnahmegeometrie mathematisch modellieren zu können) zusätzlich zum Luftbild, dessen innere und äußere Orientierung sowie ein DOM benötigt.

Rechnerisch erfolgt sie nach folgendem Prinzip [Kraus, 2004]:

1.) In der *XY*-Ebene des Orthophotos wird eine Bildmatrix ausgebreitet. Die gewählte Rasterweite entspricht der Auflösung des Orthophotos.

2.) Für jeden Rasterpunkt des Orthophotos wird ein Höhenwert *Z* aus dem DOM interpoliert *(X, Y -> X, Y, Z)*.

3.) Mithilfe der Kollinearitätsgleichungen (siehe Formel (3.1)) mit den bekannten Elementen der inneren und äußeren Orientierung des Luftbildes wird für jeden Rasterpunkt des Orthophotos *(X, Y)* ein zugehöriger Punkt (ξ, η) im Luftbild bestimmt $(X, Y, Z \rightarrow \xi, \eta)$:

4.) Da ein Pixelmittelpunkt des Luftbildes nicht exakt auf einen Pixelmittelpunkt des Orthophotos abgebildet wird, werden die Grauwerte für die Bildpunkte des Orthophotos interpoliert.

Das Prinzip der Entzerrung ist vereinfacht in Abbildung 3.2 dargestellt.

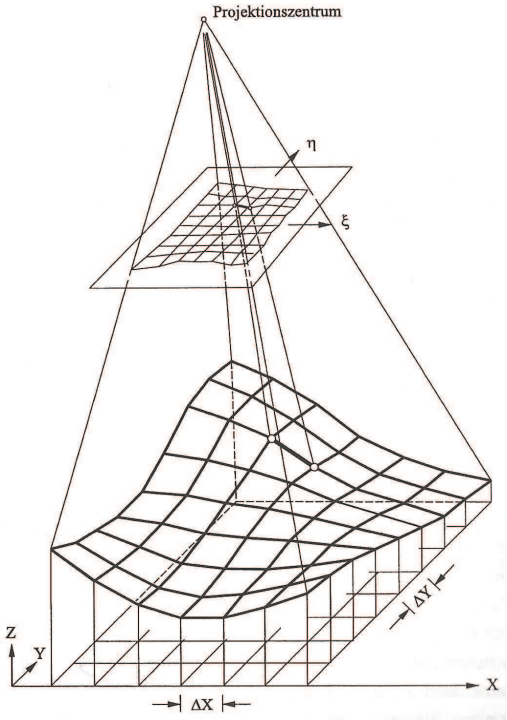

Abbildung 3.2: Zusammenhang von Orthophoto (X, Y) und verzerrtem Luftbild (ξ, η) [Kraus, 2004].

3.3.1 Genauigkeit des Orthophotos

Die Genauigkeit eines Orthophotos wird von verschiedenen Parametern beeinflusst. Da Orthophotos Folgeprodukte sind, welche auf den zu entzerrenden Photos und einem DOM basieren, hängt ihre Qualität zu einem großen Teil von der Qualität dieser Daten ab, im Speziellen von:

- der Qualität und Auflösung der zu entzerrenden Photos
- der inneren und äußeren Orientierung der Photos
- der Genauigkeit des DOMs

Im Rahmen dieser Arbeit liegt der Fokus auf der Genauigkeit des DOMs und es werden Ansätze zur Steigerung dieser untersucht. Inwiefern sich die Genauigkeit des DOMs auf die des Orthophotos auswirken kann, wird im folgenden Unterkapitel erläutert.

3.3.1.1 Reliefverschiebung

Für alle Bildbereiche außerhalb des Nadirs kommt es im Luftbild zu einer Reliefverschiebung. Ursache dafür ist die zentralperspektivische Darstellung des optischen Systems einer Kamera. Objekte, die sich direkt unter dem Mittelpunkt der Kamera befinden, werden nur mit ihrer Oberfläche auf dem Luftbild dargestellt. Weiter außen liegende Objekte werden so abgebildet, dass man sowohl ihre Oberfläche als auch ihre Seiten erkennen kann [Heidelmeyer, 2008].

Höhe 1500m (Normalobjektiv) Höhe 750m (Weitwinkelobjektiv)

Abbildung 3.3: Zwei Photos aufgezeichnet von etwa der gleichen Position, aber aus unterschiedlicher Höhe. Die Reliefverschiebungen sind wesentlich kleiner wenn die Flughöhe größer ist [Nielsen, 2004].

Fehler im DHM führen zu horizontalen Fehlern im Orthophoto, da bei der Entzerrung des Luftbilds der Bildinhalt nicht korrekt auf das Orthophoto projiziert wird und machen sich in Form der zuvor beschriebenen Reliefverschiebungen bemerkbar (siehe Abbildung 3.4). Die Reliefverschiebung verstärkt sich mit zunehmendem Blickwinkel in Bezug auf die Nadirrichtung und mit der Objekthöhe. Sie ist ebenfalls von der Flughöhe und der Kamerakonstante abhängig. Je größer die beiden Faktoren sind, desto geringer ist die Verschiebung. Somit lassen sich folgende Relationen (Formel (3.2)) und durch Herausheben die Formel für den Lagefehler (Formel (3.3)) ableiten:

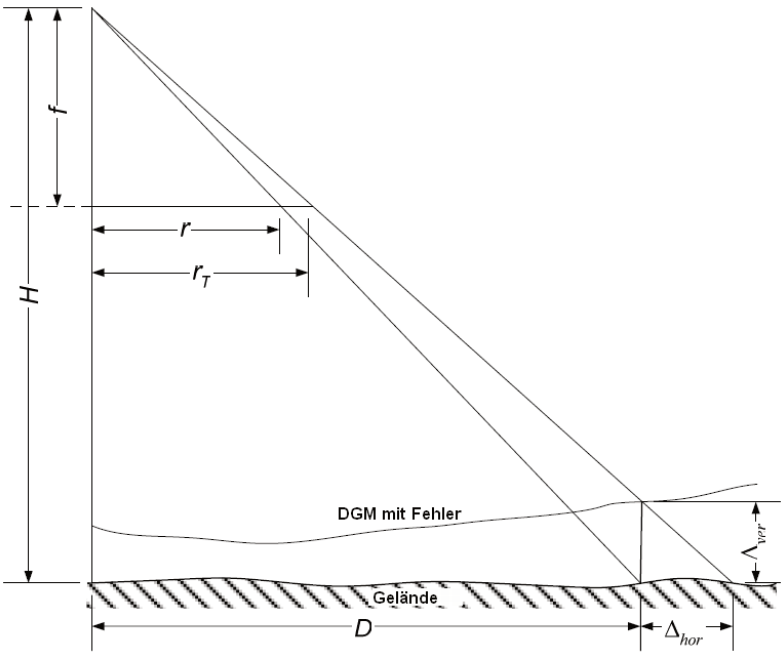

Abbildung 3.4: Vertikaler Fehler im DGM führt zu einem horizontalen Fehler im OP, der sich als Reliefverschiebung bemerkbar macht [Nielsen, 2004].

$$\frac{f}{r_T} = \frac{H}{D + \Delta_{hor}} = \frac{H - \Delta_{ver}}{D} \tag{3.2}$$

Herausheben von Δ_{hor} ergibt:

$$\Delta_{hor} = \frac{r_T \cdot \Delta_{ver}}{f} \tag{3.3}$$

mit	f	… Kamerakonstante
	H	… Flughöhe
	r_T	… Entfernung des abzubildenden Punktes vom Nadir im Bild
	D + Δ_{hor}	… Entfernung des korrekt abgebildeten Punktes vom Nadir im Gelände
	D	… Entfernung des falsch abgebildeten Punktes vom Nadir im Gelände
	Δ_{ver}	… vertikaler Fehler im DHM
	Δ_{hor}	… horizontaler Fehler

Um solche Lagefehler zu vermeiden bzw. gering zu halten, empfiehlt es sich daher, nur den zentralen Bereich von Luftbildern für die Berechnung von Orthophotos zu verwenden.

3.4 Matching von Liniensegmenten

Unter dem Matching von Liniensegmenten versteht man die Zuordnung von korrespondierenden Liniensegmenten in überlappenden Bildern (siehe Abbildung 3.5). Ein Liniensegment ist definiert über zwei Endpunkte und stellt die kürzeste Verbindung zwischen diesen dar.

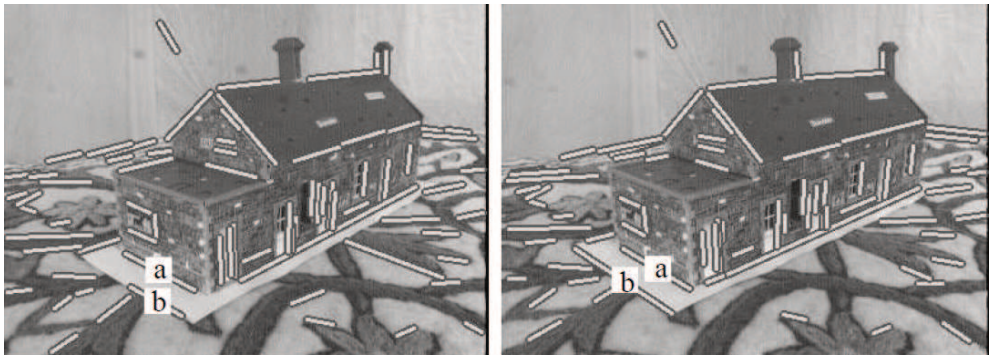

Abbildung 3.5: Matching von Liniensegmenten [Schmid, Zisserman, 1997].

Die Detektion der Liniensegmente erfolgt üblicherweise mit einem Kantendetektor, wobei der von [Canny, 1986] sehr weit verbreitet ist, aber auch andere Algorithmen, wie beispielsweise der Line Segment Detector [Von Gioi et al., 2012] für viele Anwendungsfälle gute Ergebnisse liefern. Die detektierten Liniensegmente enthalten nicht nur spektrale sondern auch wichtige geometrische und topologische Information über die aufgezeichnete Szene und die darin vorkommenden Objekte.

Die verlässliche Extrahierung von korrespondierenden Liniensegmenten in überlappenden Bildern ist für viele verschiedene Zwecke wie 3D-Objekt-Rekonstruktion, Registrierung von Photos und automatische Aerotriangulation unerlässlich. Daher stellt das Matching von Liniensegmenten eine fundamentale Aufgabe vor allem in der Photogrammetrie und Computer Vision dar und wurde in den letzten Jahrzehnten ausgiebig untersucht.

3.4.1 Vorgangsweise

Um von einem Liniensegment weitere korrespondierende Abbildungen zu finden, benötigt man charakteristische Information von diesem. Sogenannte Deskriptoren werden verwendet, um diese Information zu extrahieren und die charakteristischen Eigenschaften der Liniensegmente zu beschreiben. Dabei wird auch oft die lokale (oder auch die globale) Nachbarschaft dieser Liniensegmente herangezogen. Die Deskriptoren ermöglichen es die Ähnlichkeit von Liniensegmenten zu messen und so Korrespondenzen zu finden. Entscheidend für den Erfolg eines Deskriptors ist, dass er Liniensegmente so charakteristisch beschreibt, dass sie einerseits von anderen unterschieden und andererseits auch in einem anderen Photo (aufgezeichnet unter einer anderen Aufnahmesituation) wiedergefunden werden können. Wie man schon erahnen kann, gibt es keinen Deskriptor, der dies in allen Fällen gewährleistet. Stattdessen hängt es stark von der entsprechenden Aufnahmekonfiguration ab, welcher Deskriptor für den jeweiligen Anwendungsfall geeignet ist. In der Literatur findet man Deskriptoren, die folgende Information über das Liniensegment enthalten:

Geometrische:

- Orientierung
- Länge
- Ausmaß der Überschneidung
- Epipolargeometrie

Spektrale:

- Farbe in der Nachbarschaft
- Intensität in der Nachbarschaft
- Gradient in der Nachbarschaft

3.4.2 Herausforderungen

Das Matching von Liniensegmenten (bzw. die Definition von robusten/invarianten Deskriptoren) ist aus mehreren Gründen eine anspruchsvolle Aufgabe. Es setzt voraus, dass die Liniensegmente vorher detektiert wurden. Die Detektion ist aber oft nicht zuverlässig und weist im Allgemeinen folgende Schwächen auf [Fan, Wu, Hu, 2010], [Schmid, Zisserman, 1997]:

- die Position der Endpunkte der Liniensegmente ist unzuverlässig.
  Das hat zur Folge, dass (i) vorhandene topologische Verbindungen (zum Beispiel zwei Liniensegmente schneiden sich) verloren gehen können und (ii) Punkt-zu-Punkt-Korrespondenzen im Allgemeinen auch bei korrespondierenden Liniensegmenten unbekannt sind. Dadurch hat man bei diesen keine starken geometrischen Bedingungen wie beispielsweise die Epipolarbedingung bei korrespondierenden Punkten.

- die Fragmentierung von Liniensegmenten. So kann es vorkommen, dass auf einem einzelnen, geschlossenen Liniensegment mehrere Liniensegment–Fragmente detektiert werden. Somit können mit einer bijektiven Abbildung nicht alle Liniensegmente korrekt zugeordnet werden, was die Komplexität des Matchings weiter erhöht.

Resultierend aus den unterschiedlichen Aufnahmesituationen der Photos, vor allem in Bezug auf die Aufnahmegeometrie, gibt es weitere Herausforderungen bzw. Schwierigkeiten, die es zu berücksichtigen gilt:

- korrespondierende Liniensegmente können unterschiedliche Nachbarschaften unter anderem in Bezug auf Form, Orientierung, Farbe oder spektrale Intensität aufweisen.

Um diese Herausforderungen zu bewältigen und trotz dieser Erschwernisse zufriedenstellende Matching-Ergebnisse zu erzielen, wurden einige Ansätze entwickelt. Um einen Überblick zu bekommen, erfolgt im folgenden Kapitel eine Einteilung solcher Matching-Ansätze.

3.4.3 Einteilung von Ansätzen

Existierende Ansätze zum Matching von Liniensegmenten lassen sich grob in zwei Kategorien einteilen:

- **Liniensegmente werden einzeln gematcht**:

  Dabei werden üblicherweise Deskriptoren verwendet (siehe Kapitel 3.4.1), die vor allem geometrische und spektrale Information über die Liniensegmente und ihre Nachbarschaft enthalten und mit denen sich die Ähnlichkeit von jeweils einem Paar von Liniensegmenten berechnen lässt (beispielsweise [Wang, Wu, Hu, 2009]).

- **Liniensegmente werden in Gruppen gematcht**:

  Dies hat den Vorteil, dass mehr geometrische Information zur Verfügung steht. So können unter anderem Kriterien basierend auf den topologischen Beziehungen miteinbezogen werden wie beispielsweise bei [Wang, Neumann, You, 2009] für jeweils zwei orientierte Liniensegmente folgende:

  - Lage zueinander
  - Schnittpunkt
  - Winkel
  - Längenverhältnis

  Allerdings sind diese Ansätze im Allgemeinen wesentlich komplexer und rechenintensiver.

Des Weiteren lassen sich die Ansätze nach ihrer Eignung für bestimmte Anwendungsfälle bzw. Aufnahmekonfigurationen unterscheiden:

- **Anzahl der Bilder:**
  Die meisten Ansätze sind darauf ausgelegt jeweils zwei Bilder miteinander zu matchen. Es gibt aber auch welche, die die Integration von drei oder mehr Bildern erlauben, wodurch man stärkere geometrische Bedingungen erhält.

- **Aufnahmegeometrie:**
  Es lassen sich Anwendungsfälle, bei denen die **Basislinie** (siehe Abbildung 3.6a) im Vergleich zum Abstand zur Szene groß oder klein ist, unterscheiden.
  Die unterschiedliche Aufnahmegeometrie kann auch eine **Verdrehung** (siehe Abbildung 3.6b) oder unterschiedliche **Skalierung** (siehe Abbildung 3.6c) der Bilder zur Folge haben. Des Weiteren können **Verdeckungen** in Schattenbereichen sowie **Beleuchtungsunterschiede** (siehe Abbildung 3.6d) in den Bildern auftreten.

- **Szene:**
  Auch bezüglich der Eignung der aufgenommenen Szene lassen sich die Ansätze unterscheiden. In urbanen Gebieten liegen oft homogene Oberflächen mit einer kaum unterscheidbaren **Textur** (siehe Abbildung 3.6e) vor, die für die Performance vieler Ansätze einschränkend wirkt.
  Des Weiteren gibt es Ansätze, die speziell für Szenen, die hauptsächlich aus **ebenen Oberflächen** bestehen, wie es ebenfalls in urbanen Gebieten der Fall sein kann, ausgelegt sind.

Viele Ansätze benötigen außerdem noch (externe) Zusatzinformation wie beispielsweise:

- die **Orientierung** der Bilder, üblicherweise gegeben über die Fundamentalmatrix. Diese Information kann genutzt werden, um über die Epipolargeometrie Bedingungen für korrespondierende Liniensegmente einzuführen [Schmid, Zisserman, 1997].

- **Punktkorrespondenzen**, die zuvor in den Bildern bestimmt werden [Fan, Wu, Hu, 2010].

- **DOMe** [Chehata et al., 2002], [Jung, Paparoditis, 2003], [Tailandier, Deriche, 2002].

- **LIDAR**-Daten [Habib et al., 2010].

Zu beachten ist allerdings, dass durch die Verwendung dieser Zusatzinformation die Matching-Performance in großem Maße von deren Qualität abhängig wird.

(a) Aufnahmewinkel bzw. Basislinie (b) Verdrehung

(c) Skalierung (d) Kontrastunterschiede

(e) kaum unterscheidbare Textur

Abbildung 3.6: Ansätze des Liniensegmentmatchings können nach ihrer Robustheit gegenüber Änderungen der Aufnahmesituation und nach ihrer Eignung für Szenen eingeteilt werden [Zhang, Koch, 2012].

4 Stand der Forschung

In diesem Kapitel werden bestehende Ansätze, welche sich mit dem Matching von Liniensegmenten und der anschließenden dreidimensionalen Rekonstruktion dieser Liniensegmente beschäftigen und dem aktuellen Stand der Forschung entsprechen, vorgestellt.

4.1 Matching von Liniensegmenten

4.1.1 Matching von Liniensegmenten nach Bay et al. (2005)

Eine „State-of-the-Art"-Methode zum Matching von Liniensegmenten in zwei unkalibrierten Photos mit großer Basis wurde von [Bay et al., 2005] entwickelt. Der Aufbau des Algorithmus gliedert sich in folgende Schritte:

- Farbhistogramm-basiertes Matching:

  Durch den Vergleich von Histogrammen der Farbprofile, die zu beiden Seiten der Liniensegmente berechnet werden, erhält man erste Matching-Kandidaten.

- Topologischer Filter:

  Dieser Filter wird angewandt, um falsche Zuordnungen zu entfernen.

- Bestimmung weiterer Korrespondenzen:

  Anschließend werden die geometrischen Bedingungen des topologischen Filters verwendet, um iterativ weitere Matches hinzuzufügen, die mit der topologischen Struktur der bisherigen Matches übereinstimmen.

4.1.1.1 Farbhistogramm-basiertes Matching

Als erstes Kriterium zum Finden von korrespondierenden Liniensegmenten dienen deren Farbprofile. Und zwar wird links und rechts jedes Liniensegments ein Profil (mit der Länge des entsprechenden Liniensegments) gelegt (siehe Abbildung 4.1). Die Liniensegmente werden immer so gerichtet, dass das dunklere Farbprofil rechts ist, dadurch lassen sich auch bei verdrehten Bildern die Farbprofile eines Liniensegments vergleichen. Der Normalabstand zwischen dem Liniensegment und seinem rechten bzw. linken Farbprofil beträgt dabei 1.5 Pixel.

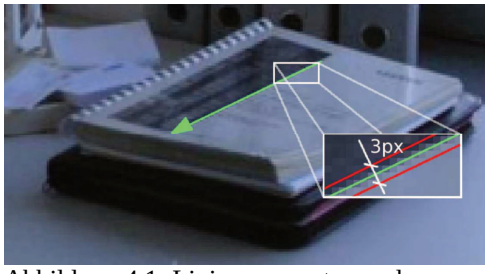

Abbildung 4.1: Liniensegmente werden so gerichtet, dass die dunklere Seite rechts ist. Der Normalabstand zwischen seinen Farbprofilen beträgt 3 Pixel [Bay et al., 2005].

Dann wird für jedes Farbprofil ein Histogramm berechnet. Dabei werden die Farbwerte diskretisiert, indem der HSV-Farbraum − mit den Achsen Hue (Farbton), Saturation (Sättigung) und Value (Hellwert), die jeweils einen Wertebereich von 0 bis 1 haben (siehe Abbildung 4.2) − in 166 Bins unterteilt wird. Konkret wird die Hue-Achse in 18 äquidistante Bereiche und die Saturation-Achse sowie die Value-Achse in 3 äquidistante Bereiche unterteilt. Zusätzlich gibt es noch 4 Grauwerte, die im Inneren des HSV-Zylinders, oder genauer gesagt in dem Bereich in, welchem die Saturation weniger als 0.25 beträgt, liegen und hier die Value-Achse in 4 Bereiche, welche die Grauwerte darstellen, unterteilen. Die weitere Unterteilung der Saturation-Achse in − wie zuvor beschrieben − 3 äquidistante Bereiche erfolgt außerhalb dieses Bereichs der Grauwerte, also im Wertebereich von 0.25 und 1. Somit ergeben sich die angesprochenen $4+18\cdot 3\cdot 3=166$ diskreten Werte.

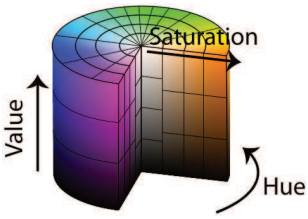

Abbildung 4.2: Zylinder-Darstellung des HSV-Farbraums mit den Achsen Hue (Farbton), Saturation (Sättigung) und Value (Hellwert) [HSV-Farbraum, 2015].

Jede Farbe des HSV-Raums kann so also eindeutig einem dieser 166 Werte zugeordnet werden. Diese Diskretisierung ist sinnvoll und notwendig, um die Farben der Profile effizient vergleichen zu können. Nach der Diskretisierung wird jedes Farbprofil noch einer Normalisierung unterzogen. Dabei geht es darum, die Profile unabhängig von der jeweiligen Länge des Profils bzw. Liniensegments zu machen, man erhält also die prozentuelle Häufigkeit, mit der jede Farbe im Profil vorkommt. Für jedes Liniensegment werden so zwei − eines für das rechte und eines für das linke Profil − normalisierte Histogramme berechnet, die anschließend verwendet werden, um die (Un-)Gleichheit zweier Liniensegmente zu berechnen. Die Ungleichheit zweier Histogramme h_1 und h_2 ist gegeben durch:

$$d_{1,2} = (\mathbf{h_1} - \mathbf{h_2})^\top \mathbf{A}(\mathbf{h_1} - \mathbf{h_2}) \tag{4.1}$$

Wobei A = $[a_{i,j}]$ eine 166 mal 166 Matrix ist, deren Elemente $a_{i,j}$ den Abstand zweier Farben $p_i = [h_i, s_i, v_i]$ und $p_j = [h_j, s_j, v_j]$ im Farbraum definieren durch:

$$\begin{aligned} a_{i,j} = & \frac{1}{\sqrt{2}} \{ (v_i s_i \cos(h_i) - v_j s_j \cos(h_j))^2 + \\ & (v_i s_i \sin(h_i) - v_j s_j \sin(h_j))^2 + (v_i - v_j)^2 \}^{1/2} \end{aligned} \tag{4.2}$$

Dadurch lässt sich ein Maß der Ungleichheit zweier Liniensegmente durch den Vergleich der rechten und linken Farbprofile mit folgender Formel angeben:

$$\sqrt{\frac{d_{1,2L}{}^2 + d_{1,2R}{}^2}{2}} \tag{4.3}$$

Mit diesem Maß werden nun alle Liniensegmente miteinander verglichen und es werden jene als Matching-Kandidaten angesehen, bei denen das Maß der Ungleichheit kleiner als 0.25 ist.

4.1.1.2 Topologischer Filter

Als Nächstes wird ein topologischer Filter angewandt, um falsche Zuordnungen zu entfernen. Dabei werden für Kombinationen von Matching-Kandidaten zwei Bedingungen (siehe Abbildung 4.3) überprüft und die relative Häufigkeit an Verstößen gegen diese Bedingungen für das jeweilige Liniensegment berechnet. Matching-Kandidaten, bei denen die relative Häufigkeit an Verstößen V mehr als 0.15 mal der relativen Häufigkeit an Verstößen V_{Max} von jener Kombination, bei der sie am höchsten ist, beträgt, werden in einem iterativen Prozess entfernt. Die zwei Bedingungen sind:

1.) Für jeweils 2 Liniensegmente und deren Matching-Kandidaten im anderen Bild:
Der Mittelpunkt von Liniensegment 1, muss in beiden Bildern auf derselben Seite des (gerichteten) Liniensegments 2 liegen.

2.) Für jeweils 3 Liniensegmente und deren Matching-Kandidaten im anderen Bild:
Der Mittelpunkt von Liniensegment 1, muss in beiden Bildern auf derselben Seite der Linie, die vom Mittelpunkt des Liniensegments 2 zum Mittelpunkt des Liniensegments 3 geht, liegen.

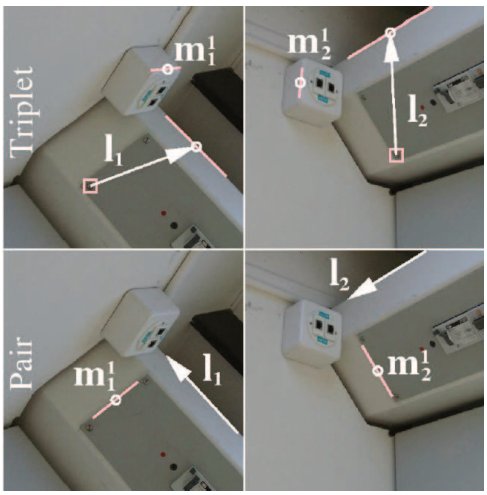

Abbildung 4.3: Geometrische Bedingungen für 2 (unten) bzw. 3 Liniensegmente (oben) und deren Matching-Kandidaten im anderen Bild [Bay et al., 2005].

Beide Bedingungen müssen theoretisch gesehen immer erfüllt sein, wenn die Liniensegmente koplanar sind und in den allermeisten Fällen, wenn sie nicht koplanar sind.

Diese Bedingungen wurden bei der vorliegenden Methode gewählt, weil der Ansatz für unkalibrierte Bilder, die nicht orientiert sind und für die daher keine Fundamentalmatrix gegeben ist, ausgelegt ist. Da sich die Fundamentalmatrix nicht direkt mithilfe der Liniensegment-Matches berechnen lässt, wird die Epipolarbedingung bei diesem Ansatz nicht als Filter eingesetzt wie bei der Methode von Schmid und Zisserman (siehe Kapitel 4.1.2), welche kalibrierte Bilder voraussetzt.

Nach Anwendung des topologischen Filters wird für jedes Liniensegment nur der plausibelste Matching-Kandidat behalten, dies ist jener mit der geringsten Summe von Ungleichheits-Maß und relativer Häufigkeit an Verstößen.

4.1.1.3 Bestimmung weiterer Korrespondenzen

Nachdem nun erste Korrespondenzen gefunden und falsche Zuordnungen mithilfe des topologischen Filters entfernt wurden, wird dieser nun verwendet, um iterativ weitere Matches hinzuzufügen, die mit der topologischen Struktur der bisherigen Korrespondenzen übereinstimmen. Dabei werden die geometrischen Bedingungen des topologischen Filters für jede Kombination von ungematchten Liniensegmenten aus den beiden Bildern, zusammen mit allen bereits gematchten Liniensegmenten, überprüft. Es gelten die gleichen Kriterien wie bisher, also Ungleichheits-Maß < 0.25 (siehe Kapitel 4.1.1.1) und $V \leq 0.15 \cdot V_{Max}$ (siehe Kapitel 4.1.1.2). Sind diese erfüllt, werden die Liniensegmente einander zugeordnet. Auch hier wird für jedes Liniensegment letztendlich nur der plausibelste Matching-Kandidat behalten. Durch diese iterative Vorgangsweise können in diesem Schritt weitere korrespondierende Liniensegmente gefunden werden, die fälschlicherweise durch den topologischen Filter entfernt oder aufgrund der hohen Ähnlichkeit mit einem anderen Liniensegment einander nicht zugeordnet wurden.

4.1.2 Matching von Liniensegmenten nach Schmid und Zisserman (1997)

Eine weitere „State-of-the-Art"-Methode zum Matching von Liniensegmenten wurde von [Schmid, Zisserman, 1997] entwickelt. Diese ist allerdings im Gegensatz zum Ansatz von [Bay et al., 2005] nur für kalibrierte Photos geeignet, sie setzt also die relative Orientierung der Bilder in Form der Fundamentalmatrix voraus. Schmid und Zisserman haben sowohl einen Ansatz für Photos mit kurzer, als auch einen für Photos mit langer Basis entwickelt. Außerdem haben sie eine Erweiterung entworfen, mit welcher das Matching auch für 3 Bilder angewandt werden kann und bei welcher nicht die Fundamentalmatrix, sondern der Trifokale Tensor vorausgesetzt wird. Ein Nachteil gegenüber dem Ansatz von [Bay et al., 2005] ist, dass die Orientierung vorausgesetzt wird. Dies kann aber genauso auch als Vorteil gesehen werden, da mithilfe der gegebenen Orientierung zusätzliche geometrische Bedingungen zwischen korrespondierenden Liniensegmenten über die Epipolargeometrie definiert werden. Die Methode ist vor allem für Bilder geeignet, die großteils ebene Oberflächen abbilden, wie es oft im urbanen Raum der Fall ist. Es ist nämlich eine weitere Voraussetzung, dass sich die Szene lokal durch eine Ebene oder durch Verschneidungen von Ebenen approximieren lässt.

4.1.2.1 Deskriptor

Der Deskriptor, sprich das Maß für die Gleichheit zweier Liniensegmente, ist für die beiden Ansätze mit kurzer bzw. langer Basis gleich. Er ist definiert über den Durchschnitt der Kreuzkorrelationen der Intensitäten in den Nachbarschaften korrespondierender Punkte dieser Liniensegmente. Der Unterschied der beiden Ansätze liegt in der Bestimmung der korrespondierenden Punkte.

4.1.2.2 Ansatz für kurze Basis

Dieser Ansatz hat das Ziel, korrespondierende Punkte direkt auf den Liniensegmenten zu finden und dann mit dem zuvor beschriebenen Deskriptor die Gleichheit zweier Liniensegmente über die Intensität in den Nachbarschaften dieser Punkte zu berechnen. Zum Finden der Korrespondenzen werden geometrische Beziehungen aus der Epipolargeometrie (siehe Abbildung 4.4) herangezogen. Diese setzen die relative Orientierung zwischen den Bildern in Form der Fundamentalmatrix F voraus. Korrespondierende Punkte (x in Bild 1, x' in Bild 2) müssen die Epipolarbedingung erfüllen:

$$x'^T Fx = 0 \tag{4.4}$$

Ein Punkt x in Bild 1 definiert eine Epipolarlinie l_e' in Bild 2, auf welcher sich der korrespondierende Punkt x' befinden muss (siehe Abbildung 4.4):

$$l_e' = Fx \quad \text{bzw.} \quad l_e = F^T x' \tag{4.5}$$

Mit der Epipolargeometrie lassen sich für korrespondierende Liniensegmente l und l' Punkt-zu-Punkt-Korrespondenzen generieren. Da der korrespondierende Punkt x' sowohl am Liniensegment l' als auch auf der Epipolarlinie l_e' liegt, findet man ihn über den Schnittpunkt dieser beiden, sofern l' keine Epipolarlinie ist (siehe Punkt y' in Abbildung 4.6):

$$x' = l'x l_e' = l'xFx \tag{4.6}$$

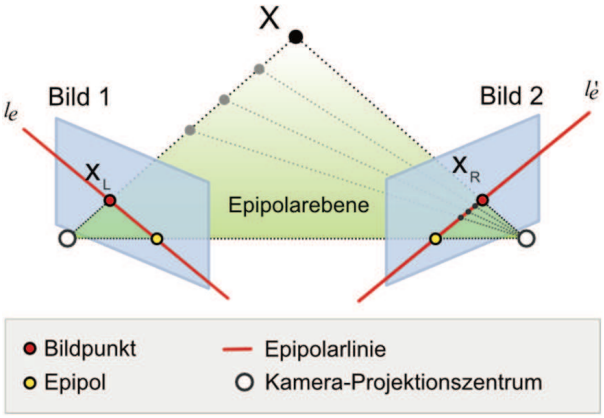

Abbildung 4.4: Begriffe und geometrische Beziehungen aus der Epipolargeometrie [Epipolargeometrie, 2015].

Für jedes Liniensegment in Bild 1 wird die Gleichheit zu allen Liniensegmenten in Bild 2 innerhalb eines Suchraums (dieser wird im nächsten Absatz beschrieben) berechnet, indem diese als korrespondierend angenommen werden und so Punkt-zu-Punkt-Korrespondenzen generiert werden. Es wird dann aber nur jenes Liniensegmente-Paar mit der größten Gleichheit als korrekte Zuordnung behalten.

Der Suchraum ist über den Epipolarstrahl definiert, welcher aus den Epipolarlinien in Bild 2 der Endpunkte des Liniensegments in Bild 1 besteht, und das korrespondierende Liniensegment zumindest teilweise enthalten muss.

Dieser Ansatz, der auf der Bestimmung der Kreuzkorrelation beruht, versagt allerdings, wenn korrespondierende Nachbarschaften aufgrund der verschiedenen Aufnahmegeometrie der beiden Bilder signifikante Unterschiede aufweisen. Für solche Fälle wurde ein Ansatz entwickelt, der für Bilder mit langer Basis geeignet ist und im Folgenden beschrieben wird.

4.1.2.3 Ansatz für lange Basis

Durch die bekannte relative Orientierung zweier Bilder − gegeben durch die Fundamentalmatrix – lässt sich durch korrespondierende Bildinhalte die 3D-Struktur der Szene bis auf eine projektive Ambiguität des 3D-Raums rekonstruieren. So bestimmen 2 korrespondierende Linien eine Linie im 3D-Raum und umgekehrt legt eine Linie im 3D-Raum ein Ebenenbüschel und somit eine Homographie fest (siehe Abbildung 4.5).

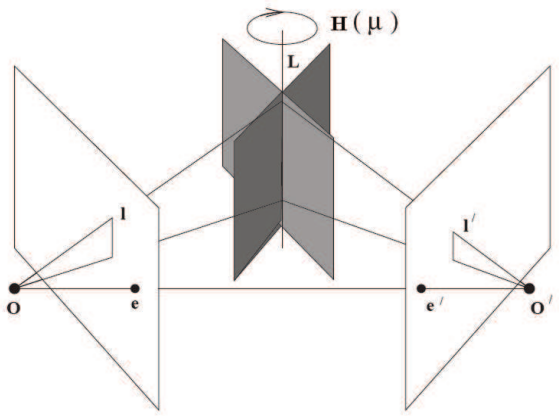

Abbildung 4.5: Die Linien l und l' bestimmen eine 3D-Linie L, welche ein Ebenenbüschel und somit eine Homographie H(μ) festlegt [Schmid, Zisserman, 1997].

Da Liniensegmente häufig an den Übergängen zweier Ebenen, sprich auf Kanten, liegen, kann die Homographie in der Nachbarschaft beider Seiten des Liniensegments unterschiedlich sein. Daher wird bei diesem Ansatz die Gleichheit von Liniensegmenten für beide Seiten der Liniensegmente separat berechnet und zwar jeweils in deren Nachbarschaft für einen rechteckigen Streifen mit einer Breite von 14 Pixel. Als Kriterium für das Matching wird letztendlich aber nur jene Seite mit der höheren Korrelation herangezogen. Um korrespondierende Punkte in den Nachbarschaften der Liniensegmente zu finden und somit die Gleichheit zweier Liniensegmente berechnen zu können, muss jedes Mal der Parameter μ der Homographie $H(\mu)$ festgelegt werden (siehe Abbildung 4.6). Der Wert μ^* für den Parameter μ wird bestimmt, indem man für den Punkt x aus Bild 1 den korrespondierenden Punkt $x'(\mu^*)$ in Bild 2 entlang der Epipolarlinie Fx sucht. Dabei werden 10 verschiedene Punkte $x'(\mu)$ in äquidistanten Abständen entlang der Epipolarlinie und innerhalb eines festgelegten Abstands zum Liniensegment l' verglichen und jener Punkt $x'(\mu^*)$, bei dem die Kreuzkorrelation mit dem Punkt x am größten ist, als der korrespondierende angenommen. Der Punkt x wird dabei in der Nachbarschaft des Liniensegments l und zwar im Eckpunkt des zuvor beschriebenen rechteckigen Streifens gewählt.

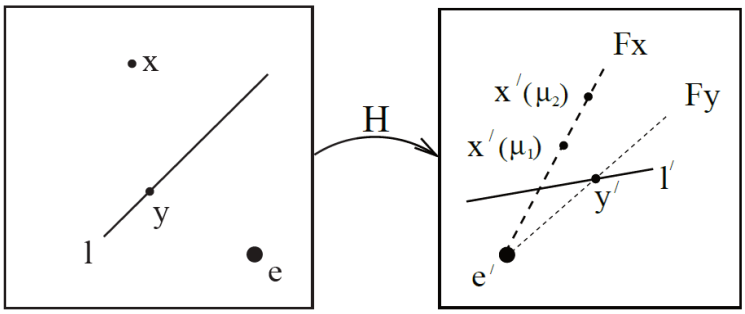

Abbildung 4.6: Wenn μ variiert, wird der Punkt x, welcher nicht auf der Linie l liegt, von $H(\mu)$ auf einen Punkt $x'(\mu)$ abgebildet, welcher sich entlang der Epipolarlinie Fx bewegt. Der Punkt y, welcher auf der Linie l liegt, wird bei allen Werten von y auf einen fixen Punkt y' abgebildet [Schmid, Zisserman, 1997].

4.1.2.4 Erweiterung auf drei Bilder

Durch die Erweiterung des Ansatzes auf 3 Bilder lassen sich zusätzliche geometrische Bedingungen einführen, mit denen die beim Ansatz mit kurzer oder langer Basis gefundenen Korrespondenzen verifiziert werden können. Die Trifokalgeometrie liefert als Erweiterung der Epipolargeometrie geometrische Bedingungen für korrespondierende Linien in drei Bildern (siehe Abbildung 4.7).

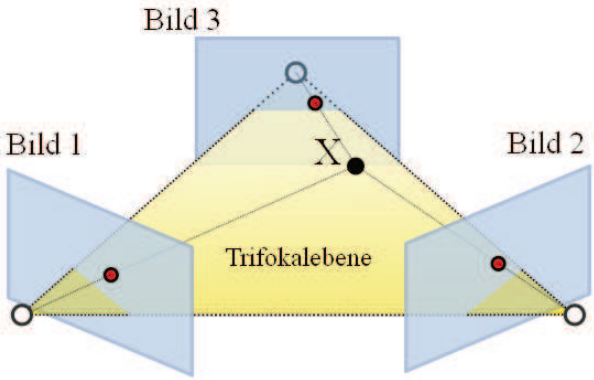

Abbildung 4.7: Schema der Trifokalgeometrie [Trifokalgeometrie, 2015].

Ist die Position eines Objektpunktes in zwei Bildern bekannt, so ist seine Position im dritten Bild der Schnittpunkt der beiden Epipolarlinien. Somit lässt sich über diese Trifokalbedingung auch die Position eines Liniensegments, welches in 2 Bildern bekannt ist, im dritten Bild vorhersagen. Die Überprüfung, ob diese Trifokalbedingung genau genug eingehalten wird, ist der erste Teil dieser Verifizierung. Der zweite Teil ist eine zusätzliche Kontrolle der Ähnlichkeit der Intensität in der Nachbarschaft der Liniensegmente. Durch diese beiden Schritte lassen sich laut [Schmid, Zisserman, 1997] in der Praxis alle Fehlzuordnungen eliminieren.

4.1.3 Implementierung zum Matching von Liniensegmenten und zur Rekonstruktion von 3D-Kanten von Werner und Zisserman (2002)

Eine Matlab-Implementierung zum Matching von Liniensegmenten und zur anschließenden Rekonstruktion der korrespondierenden Liniensegmente zu 3D-Kanten wurde 2002 von [Werner, Zisserman, 2002] umgesetzt. Der zugrunde liegende Ansatz basiert großteils auf der Matching-Methode von [Schmid, Zisserman, 1997], wie sie in Kapitel 4.1.2 beschrieben wird, erweitert diese aber auch auf mehr als drei Bilder. Der Input für den Algorithmus ist ein Satz von Grauwert-Bildern, jeweils mit einer Projektionsmatrix. Optional kann auch ein Satz von Liniensegmenten für jedes Bild angegeben werden, ansonsten werden diese mit dem Canny-Kantendetektor [Canny, 1986] bestimmt. Als Output erhält man einen Satz von korrespondierenden Liniensegmenten sowie eine 3D-Rekonstruktion dieser Liniensegmente. Die Rekonstruktion der 3D-Liniensegmente erfolgt unter Verwendung einer nicht-linearen Methode, der Maximum-Likelihood-Schätzung. Der Algorithmus zum Matching der Liniensegmente in zwei Schritten wird im Folgenden beschrieben, wobei die fett gedruckten Begriffe im Anschluss genauer erläutert werden.

4.1.3.1 Algorithmus

1.) Bestimmung vorläufiger Matches

Bestimme einen Satz von Bild-Paaren mit großer Überschneidung.
Für jedes Bild-Paar führe Folgendes aus:

a) Füge alle Matches zur Liste von Matches ζ hinzu. Ein Match S liegt dann vor, wenn die **Epipolarstrahlbedingung** erfüllt ist und der **Gleichheitswert** c größer als 0.6 ist.

b) Für jeden Match S suche nach weiteren korrespondierenden Segmenten beginnend bei dem Bild mit der größten Überschneidung. Füge das Liniensegment s zum Match S hinzu, wenn die **Rückprojektionsbedingung** erfüllt ist und der **Gleichheitswert** c größer als 0.6 ist.

2.) Bestimmung finaler Matches

Wähle die Teilmenge ζ^* von ζ, bei der alle Paare von korrespondierenden Liniensegmenten konsistent und bei der die Gleichheitswerte maximal sind, wobei ein korrespondierendes Liniensegmente-Paar dann konsistent ist, wenn entweder die **Einzigartigkeitsbedingung** oder die **Ordnungsbedingung** erfüllt ist.

Gleichheitswert:

Der Gleichheitswert c ist im Prinzip der Deskriptor, der ein Maß für die Ähnlichkeit zweier Liniensegmente s, s' darstellt. Er wird folgendermaßen berechnet:

1.) Für jeden Punkt p von s führe Folgendes aus:

a) Erzeuge für s und s' jeweils 3 Fenster. Eines zentriert in p, eines links von p und eines rechts von p.

b) Berechne die normalisierte Kreuzkorrelation der Fenster von s mit den zugehörigen Fenstern von s’, welche über die Homographie bestimmt werden, wie beim Ansatz von Schmid und Zisserman für lange Basis (siehe Kapitel 4.1.2.3).
c) Nimm als Kriterium nur jenes Fenster-Paar mit der höchsten Korrelation.

2.) Wenn es mehr als 10 Punkte p mit einer Korrelation c_i > 0.6 gibt, dann ist der Gleichheitswert c für die Liniensegmente s und s' gleich dem Durchschnitt von allen c_i die größer als 0.6 sind, ansonsten ist c gleich 0.

Epipolarstrahlbedingung:

Die Epipolarstrahlbedingung muss für korrespondierende Liniensegmente s in Bild 1 und s' in Bild 2 gelten und wird daher beim Matching überprüft. Sie besteht aus folgenden 2 Bedingungen:

1.) Das aus s und s' mittels Vorwärtsschnitt rekonstruierte 3D-Liniensegment darf nicht mit den zugehörigen Projektionszentren in einer Ebene, der Epipolarebene, liegen. In diesem Fall wäre nämlich das Ergebnis der Kreuzkorrelation, auf welchem der Gleichheitswert beruht, ungenau.

2.) s bestimmt über den Epipolarstrahl einen Suchraum, welcher aus den Epipolarlinien in Bild 2, die von den Endpunkten von s definiert werden, besteht. s' muss in diesem Suchraum liegen oder ihn zumindest schneiden.

Rückprojektionsbedingung:

Die Rückprojektionsbedingung wird für einen bestehenden Match S und ein Liniensegment s überprüft. Sie ist erfüllt, wenn folgende drei Bedingungen zutreffen:

1.) Die Residuen zwischen den Liniensegmenten – sowohl von S als auch s – und der Rückprojektion des 3D-Liniensegments, welche aus ebendiesen Liniensegmenten rekonstruiert wird, sind klein.

2.) Die Rückprojektion von S in das Bild des Liniensegments s überschneidet sich mit s.

3.) Die Rückprojektion von S in das Bild des Liniensegments s hat die gleiche Orientierung wie s. Die Orientierung ist dadurch festgelegt, dass die hellere Seite immer rechts ist.

Einzigartigkeitsbedingung:

Die Einzigartigkeitsbedingung ist dann erfüllt, wenn im gesamten Satz von Matches ζ jedes Liniensegment nur in einem Match S vorkommt. Eine Ausnahme kann aufgrund der Fragmentierung von Liniensegmenten, die bei der Detektion auftreten kann, gemacht werden und zwar dann, wenn die Liniensegmente anschließend zu einem zusammengefügt werden können und dann ohnehin wieder einzigartig sind.

Ordnungsbedingung:

Die **Ordnungsbedingung** wird für ein korrespondierendes Liniensegmente-Paar überprüft, wenn eines der Liniensegmente in einem anderen verifizierten Match bereits vorkommt und das durch Vorwärtsschnitt rekonstruierte 3D-Liniensegment dieses verifizierten Matches mit den zugehörigen Projektionszentren in einer Ebene, der Epipolarebene, liegt. Sie ist dann erfüllt, wenn das neue Liniensegmente-Paar auch (annähernd) in dieser Ebene liegt.

5 Daten

Luftbilder und ein Oberflächenmodell von Graz wurden von der Firma Meixner Imaging GmbH zur Verfügung gestellt. Bei dem übergeordneten Koordinatensystem, auf welches sich das Geländemodell sowie die Parameter der äußeren Orientierung der Luftbilder beziehen, handelt es sich um die UTM-Abbildung der Zone 33N, welche das Rotationsellipsoid WGS 84 verwendet.

5.1 Luftbilder

Die Befliegung über Graz wurde am 22. September 2007 mit einer Ultracam-X [Ultracam, 2015] durchgeführt, wobei die verwendeten Streifen zwischen 9 Uhr 30 und 11 Uhr (UTC) aufgezeichnet wurden. Die Überlappung der Photos beträgt durchschnittlich etwa 80% in Flugrichtung − also für benachbarte Photos eines Streifens − und etwa 60% senkrecht zur Flugrichtung − also für benachbarte Streifen. Die Photos wurden relativ senkrecht zum Gelände aufgenommen, die Stereoanordnungen entsprechen somit annähernd einem Normalfall. Die Georeferenzierung erfolgte direkt wie üblich mittels GPS und IMU. Die dadurch bestimmten Parameter der äußeren Orientierungen der Aufnahmen − die drei Rotationswinkel, welche die Verdrehung des Bildkoordinatensystems gegenüber dem übergeordneten Koordinatensystem beschreiben und die drei Koordinaten des Projektionszentrums im übergeordneten Koordinatensystem − sowie die bekannte innere Orientierung der Kamera − die Kamerakonstante und die zwei Bildkoordinaten des Hauptpunktes − sind für alle Luftbilder gegeben. Die Luftbilder liegen im TIF-Format vor. Ihre Auflösung beträgt jeweils 14430 mal 9420 Pixel, was etwa 135 Megapixel pro Kanal entspricht. Es handelt sich Echtfarbbilder, deren Kanäle Rot, Grün und Blau durch einen Pan-sharpening Prozess erhalten wurden. Die Bit-Tiefe pro Kanal wurde von ursprünglich 14 bits auf 8 bits reduziert. Die Bilder wurden aus eine Höhe von durchschnittlich etwa 1500 m Höhe über dem Meeresspiegel bzw. etwa 1100 m über dem Gelände aufgezeichnet. Aus dem Verhältnis der durchschnittlichen Höhe über dem Gelände und der bekannten Kamerakonstante von 100.5 mm ergibt sich ein durchschnittlicher Maßstab von etwa 1:10945 für die Luftbilder. Die Ausdehnung eines Pixels im Bild beträgt 0.0072 mal 0.0072 mm², was multipliziert mit dem durchschnittlichen Maßstab eine durchschnittliche Ausdehnung eines Pixels im Gelände von etwa 80 mal 80 mm² ergibt.

5.2 Digitales Oberflächenmodell

Das gegebene DOM ist unterteilt in Kacheln mit jeweils einer Auflösung von 4352 mal 4352 Pixel, wobei jedes Pixel eine Ausdehnung von 100 mal 100 mm² hat. Das DOM liegt im TIF-Format vor, es beinhaltet einen Kanal mit einer Bildtiefe von 32 Bit, welcher die Höhe der Oberfläche über dem Meeresspiegel angibt und als Fließkommazahl verspeichert ist. Der Wertebereich, also die Höhe über dem Meeresspiegel, liegt für das untersuchte Gebiet zwischen 300 und 450 m.

6 Methoden

Zur genauen Modellierung von Traufenkanten wurde im Zuge dieser Masterarbeit ein Ansatz entwickelt, der sich in folgende Schritte gliedert:

1.) Detektion von Liniensegmenten in allen Luftbildern mit anschließendem Clustering.

2.) Matching der Liniensegmente.

3.) Rekonstruktion von 3D-Kanten durch Vorwärtsschnitt von korrespondierenden Liniensegmenten. Gegebenenfalls Ausgleichung bei überbestimmten Vorwärtsschnitt.

6.1 Detektion von Liniensegmenten

Zur Detektion von Liniensegmenten gibt es einige Ansätze, die sich bewährt haben. Der Kantendetektor von [Canny, 1986] ist sehr weit verbreitet, aber auch andere Algorithmen wie der Line Segment Detector (LSD) [Von Gioi et al., 2012] liefern in vielen Anwendungsfällen gute Ergebnisse. Da eine bestehende Implementierung des LSD-Algorithmus zur Verfügung stand, wurde dieser zur Detektion der Liniensegmente verwendet. Der LSD-Algorithmus detektiert Liniensegmente in Subpixelgenauigkeit. Er kann auf allen digitalen Bildern angewandt werden und benötigt keine Parameter-Einstellungen des Nutzers. Im Durchschnitt wird ein falsches Liniensegment pro Bild detektiert. Die in den gegebenen Luftbildern mit dem LSD-Algorithmus detektierten Liniensegmente sind in Abbildung 6.1 in der Mitte ersichtlich. Eine Abschätzung der Genauigkeit der detektierten Liniensegmente ist in Kapitel 6.3.2 über σ_0 gegeben.

6.1.1 Clustering der Liniensegmente

Es ist sinnvoll, das Ergebnis der Liniensegment-Detektion einem Clustering zu unterziehen, da es immer wieder vorkommt, dass entlang einer Kante mehrere (kurze) Liniensegmente detektiert werden. Ziel des Clusterings ist es, in solchen Fällen die (kurzen) Liniensegmente zu einem längeren zusammenzufassen bzw. zu verbinden. Hierzu wurde ein möglichst simpler und einfach zu implementierender Ansatz gewählt, der aber dennoch seinen Zweck erfüllt und nun beschrieben wird.

An den Endpunkten aller Liniensegmente wird im Umkreis von **10 Pixel** nach dem nächsten Endpunkt gesucht, welcher (i) einem anderen Liniensegment angehört, (ii) dessen Liniensegment eine ähnliche Richtung hat (der Winkel darf sich um maximal **0.5°** unterscheiden) und (iii) dessen Normalabstand maximal **2 Pixel** beträgt. Die gewählten Werte für die Parameter haben sich bei einer empirischen Untersuchung mit den vorliegenden Daten als sinnvoll erwiesen. Das Resultat nach dem Clustering ist in Abbildung 6.1 rechts dargestellt.

Abbildung 6.1: Ausschnitt eines Luftbilds ohne Liniensegmente (links), mit detektierten Liniensegmenten (Mitte) und mit detektierten, geclusterten Liniensegmenten (rechts).

6.2 Matching von Liniensegmenten

Zum Matching von Liniensegmenten wurde ein Ansatz entwickelt, der einerseits auf den spektralen Eigenschaften in der Nachbarschaft der Liniensegmente beruht und sich dabei großteils an dem Ansatz von [Bay et al., 2005] (siehe Kapitel 4.1.1) orientiert, aber auch gewisse Unterschiede zu diesem aufweist. So werden außerdem auch Zusatzinformationen aus dem gegebenen DOM und der Orientierung der Luftbilder miteinbezogen. Der Ansatz wird im Folgenden beschrieben.

Ein Segment wird zu einem anderen als korrespondierend angenommen, wenn es ihm im spektralen Merkmalsraum am Nähesten ist und eine gewisse Maximaldistanz unterschreitet. Zudem können mithilfe der Orientierung der Luftbilder und eines groben, aber eben nicht exakten DOMs (Ziel ist es ja schließlich auch dieses DOM zu verbessern), räumliche Merkmale grob abgeschätzt werden, wodurch vorab zusätzliche Nebenbedingungen aufgestellt werden und damit die Suche nach Korrespondenzen auf einige Matching-Kandidaten eingeschränkt wird.

6.2.1 Spektraler Merkmalsraum

Im spektralen Merkmalsraum dienen die Farbprofile der Segmente als Kriterium. Da direkt auf einer Bruchkante die spektralen Eigenschaften aufgrund der diffusen Reflexionseigenschaften stark mit geometrischen Faktoren wie dem Aufnahmewinkel variieren, wird nicht direkt entlang des Liniensegments, sondern stattdessen links und rechts davon ein Profil (mit der Länge des entsprechenden Liniensegments) gelegt wie bei [Bay et al., 2005] (siehe Abbildung 3.1). Wenn das Liniensegment auf der Schnittgeraden zweier Ebenen oder am Übergang zwischen spektral unterschiedlichen Flächen liegt, so sind die Farbprofile ein geeigneter Deskriptor für das Liniensegment. Deshalb werden sie auch verwendet, um die Nähe von Liniensegmenten im spektralen Merkmalsraum zu berechnen, womit sie ein wichtiges Kriterium für das Matching sind. Um auch bei verdrehten Luftbildern die richtigen Farbprofile eines Liniensegments vergleichen zu können, werden die Liniensegmente immer so gerichtet, dass das dunklere Farbprofil rechts ist (siehe Abbildung 3.1). Da alle Luftbilder aus einem relativ senkrechten Winkel zur Erdoberfläche aufgenommen wurden, die Stereoanordnungen also nahezu einem Normalfall entsprechen, sind die zu erwartenden geometrieabhängigen Reflexionsunterschiede bei korrespondierenden Segmenten bzw. deren Farbprofilen aus verschiedenen Luftbildern gering. Zusätzlich könnte man sich auf die zentralen Bereiche der Luftbilder beschränken, um diesen Einfluss weiter zu vermindern, was in dieser Arbeit aber nicht gemacht wurde. Um die Wahrscheinlichkeit, dass Farbprofile korrespondierender Segmente möglichst dieselbe Fläche (beispielsweise bei einer Traufenkante kann man davon ausgehen, dass sich auf beiden Seiten eine Ebene befindet, auf der einen das Dach und auf der anderen Seite die Fassade bzw. der Gehsteig) abbilden, hoch zu halten, ist der Normalabstand zwischen dem Liniensegment und dessen rechtem bzw. linkem Farbprofil möglichst klein zu wählen. Wie von [Bay et al., 2005] vorgeschlagen, wurde ein Normalabstand von 1.5 Pixel gewählt. Auch eine empirische Untersuchung der Luftbilder konnte dies als sinnvollen Wert bestätigen, da in den meisten Fällen die charakteristischen Flächen zu beiden Seiten des Segments abgebildet werden und die spektrale Variabilität, welche direkt an Bruchkanten auftritt, hier kaum noch zu beobachten ist (siehe Abbildung 3.1).

Auch wenn die geometrieabhängigen Reflexionsunterschiede eher gering sind, gibt es doch Fälle, in denen sie relevant sind. So kommt es vor, dass in einem Luftbild die Fassade eines Hauses sichtbar ist und in einem anderen nicht. Von den Liniensegmenten der entsprechenden Traufenkante ist nur das dachseitige Farbprofil vergleichbar, das andere Farbprofil jedoch nicht, da es ja bei einem Segment die Fassade abbildet und bei einem anderen beispielsweise den Gehsteig (siehe Abbildung 6.2).

Abbildung 6.2: Die Fassade des Hauses ist im linken Luftbild sichtbar, im rechten hingegen nicht. Daher sind von den Farbprofilen (rot) des Liniensegments (gelb) nur jene auf der Dachfläche vergleichbar.

Dieser Spezialfall wird bei [Bay et al., 2005] nicht berücksichtigt, im vorliegenden Ansatz allerdings schon. Und zwar dadurch, dass, wenn die nach [Bay et al., 2005] berechnete Gleichheit (siehe Kapitel 4.1.1.1 bzw. Formel (4.1)) der Farbprofile auf einer Seite um einen wählbaren Faktor (empirische Untersuchungen mit den gegebenen Daten haben gezeigt, dass **15** ein sinnvoller Wert ist) größer ist, nur diese Seite für die Berechnung der Nähe im spektralen Merkmalsraum herangezogen wird. Anstelle von Formel (4.3) ist das Maß der Ungleichheit zweier Liniensegmente dann über $d_{1,2L}$ bzw. $d_{1,2R}$ aus Formel (4.1) definiert. Somit können auch Korrespondenzen bei Liniensegmenten gefunden werden, bei denen die Farbprofile nur auf einer Seite vergleichbar sind, diese müssen dann aber eine besonders große Ähnlichkeit aufweisen.

Weiters ist zu beachten, dass detektierte und geclusterte Liniensegmente aus verschiedenen Luftbildern, auch wenn sie dieselbe Kante abbilden, eine unterschiedliche Länge aufweisen bzw. einen anderen Teil dieser Kante projizieren können. Sie können trotzdem als korrespondierend angesehen werden, da beim Verschneiden der Ebenen, die aus Liniensegment und Projektionszentrum gebildet werden, die resultierende Schnittgerade nach wie vor durch die gesuchte Kante geht. Um zu zeigen, dass diese beim Matching auch zugeordnet werden können, werden nun jene Annahmen, die bezüglich der Reflexionseigenschaften von Flächen bzw. Ebenen wie beispielsweise Dächer oder Gehsteige und der daraus resultierenden und in den Luftbildern aufgezeichneten elektromagnetischen Strahlung (in den Bereichen R, G und B) getroffen wurden, erläutert:

- Die spektrale Ähnlichkeit von Liniensegmenten ist besonders groß, wenn sie denselben Bereich auf derselben Fläche abbilden. (Die Differenz resultiert nur aus den unterschiedlichen Aufnahmesituationen der Luftbilder.)

- Die spektrale Ähnlichkeit von Liniensegmenten ist groß, wenn sie einen unterschiedlichen Bereich auf derselben (mehr oder weniger homogenen) Fläche abbilden, da diese (mehr oder weniger) homogene Reflexionseigenschaften besitzt.

- Die spektrale Ähnlichkeit von Liniensegmenten ist in der Regel gering, wenn sie unterschiedliche Flächen abbilden, ausgenommen diese weisen ähnliche Eigenschaften bezüglich Material und Geometrie, beispielsweise bei der Rauhigkeit oder Neigung, auf. Diese Flächen haben dann auch homogene Reflexionseigenschaften und die jeweiligen Liniensegmente können im spektralen Merkmalsraum kaum unterschieden werden. Die falsche Zuordnung in einem solchen Fall kann aufgrund der Nebenbedingungen, die auf grob abgeschätzten räumlichen Merkmalen basieren, vermieden, aber nicht gänzlich ausgeschlossen werden.

Somit kann man erwarten, dass sich in der Regel jene Liniensegmente im spektralen Merkmalsraum näher sind, die zwar dieselbe Kante, aber einen anderen Bereich dieser Kante abbilden, als jene, die unterschiedliche Kanten abbilden, beziehungsweise bei mehreren Liniensegmenten auf einer Kante, jene mit der größeren Überdeckung. Die Wahrscheinlichkeit, dass Liniensegmente korrespondent sind, ist also einigermaßen proportional zur deren Nähe im spektralen Merkmalsraum und als Maß für die Nähe im spektralen Merkmalsraum dient die Formel (4.3) bzw. in dem zuvor beschriebenen Spezialfall die Formel (4.1).

Bei dem Ansatz von [Bay et al., 2005] werden in der Folge noch geometrische Nebenbedingungen eingeführt, welche in dieser Arbeit aber nicht umgesetzt werden. Stattdessen wird die Position und Lage der jeweiligen Liniensegmente abgeschätzt, um eigene Nebenbedingungen einzuführen.

6.2.2 Nebenbedingungen

Mithilfe der Orientierung des jeweiligen Luftbildes und eines vorläufigen, nicht exakten DOMs, lassen sich die 2D-Liniensegmente aus den entsprechenden Luftbildern auf das DOM projizieren und damit grob in das übergeordnete 3D-Koordinatensystem transformieren. Diese Näherungen der 3D-Liniensegmente sind zwar nicht genau, reichen aber aus, um Nebenbedingungen aufzustellen bzw. die Suche nach korrespondierenden Liniensegmenten auf einige Kandidaten einzuschränken. Als Nebenbedingung für die Liniensegmente bietet sich hierbei die Position, sprich die räumliche Nähe zweier Liniensegmente an.

Als weitere Nebenbedingung ist es sinnvoll, nur Liniensegmente miteinander zu matchen, wenn deren Ebenen – definiert über das Liniensegment und das Projektionszentrum des jeweiligen Luftbildes – beim Vorwärtsschnitt (siehe Kapitel 6.3) keinen schleifenden Schnitt erzeugen. Dies hätte nämlich aus numerischen Gründen Ungenauigkeiten zur Folge, die sich direkt auf die resultierende 3D-Schnittgerade auswirken würden.

6.2.2.1 Nebenbedingung 1 − innerhalb des Suchfensters

Um geeignete Werte für die Wahl der ersten Nebenbedingung zu finden, muss zunächst die zu erwartende Genauigkeit der Näherungen der 3D-Liniensegmente abgeschätzt werden. Ein häufig zu erwartender Fehler bei der Transformation in ein übergeordnetes, an die Erdoberfläche angenähertes Koordinatensystem, liegt dabei aufgrund der Ungenauigkeit des DOMs auf der Hand und soll an folgendem Beispiel beschrieben werden.

Der Sehstrahl von Projektionszentrum und Punkt des Liniensegments, welches im Luftbild beispielsweise eine Traufenkante abbildet, liegt bei der Projektion auf das DOM nicht auf dieser Kante, sondern schießt daran vorbei, sodass der Sehstrahl die Oberfläche erst auf Straßenniveau trifft (siehe Abbildung 6.5). Grund dafür ist die Abrundung der Traufenkante, eine häufig auftretende Ungenauigkeit in einem DOM (siehe Abbildung 6.3).

Abbildung 6.3: Links: Oberflächenmodell mit abgerundeten, ungenauen Traufenkanten. Rechts: Vorläufiges Orthophoto, bei dem die Auswirkungen der Fehler im Oberflächenmodell an der fehlerbehafteten Abbildung der Traufenkanten ersichtlich sind.

Der dadurch verursachte Fehler in der Höhenkoordinate des 3D-Punktes besteht aus der Höhendifferenz zwischen der Kante, auf der der Punkt eigentlich liegen sollte, und dem Niveau, auf das er aber projiziert wird.

Die Fehler in den Lagekoordinaten des 3D-Punktes, hängen nicht nur von der Objekthöhe, sondern auch vom Aufnahmewinkel des (jeweiligen Bildpunktes des) Luftbildes zur Erdoberfläche bzw. zum (annähernd darin liegenden) übergeordneten Koordinatensystem ab, und fällt im Nadirpunkt (sprich bei einer komplett senkrechten Aufnahme) gänzlich weg. Da die Luftbilder einen genäherten Normalfall bilden und der Aufnahmewinkel relativ senkrecht zur Erdoberfläche steht, ist der zu erwartende Lagefehler für diesen Fall deutlich geringer. Allerdings wird der Aufnahmewinkel vom Mittelpunkt des Luftbildes, wo er annähernd orthogonal ist, zum Bildrand hin immer größer, wodurch hier auch der Lagefehler zunimmt. Daher beschränkt man sich bei der Orthophotoherstellung oft auf die zentralen Bereiche der Luftbilder. Selbiger Fehler ist es auch, der in Orthophotos zu Reliefverschiebungen führt (siehe Kapitel 3.3.1.1) und damit ein Hauptgrund, warum das DOM verbessert werden soll, um genauere Orthophotos produzieren zu können. Somit lässt sich der maximal auftretende Lagefehler Δ_{hor} nach der Formel (3.3) durch Einsetzten der Kamerakonstante f, der Entfernung eines Bildrands im Luftbild vom Nadir r_T sowie einer maximalen Objekthöhe als der angenommene vertikale Fehler im DHM Δ_{ver} berechnen. Als maximale Objekt- bzw. Gebäudehöhe erscheint für Graz ein Wert von 30 Metern sinnvoll zu sein.

Das Einsetzen mit

	f	= 100.5 mm
	r_T	= 51.948 mm in X-Richtung bzw. 33.912 mm in Y-Richtung
	Δ_{ver}	= 30 m bzw. 30000 mm
	m_B	= 10945

ergibt nach Formel (3.3) den maximalen Lagefehler Δ_{hor}, wobei durch die durchschnittliche Maßstabszahl m_B dividiert wird, um den Fehler im Bildkoordinatensystem zu erhalten:

$$\Delta_{hor} = \frac{r_T \cdot \Delta_{ver}}{f \cdot m_B} = \frac{51.948 \cdot 30000}{100.5 \cdot 10945} = 1.417\ mm$$

bzw.

$$\Delta_{hor} = \frac{r_T \cdot \Delta_{ver}}{f \cdot m_B} = \frac{33.912 \cdot 30000}{100.5 \cdot 10945} = 0.925\ mm$$

Somit erhält man einen maximalen Lagefehler von 1.417 mm bzw. 197 Pixel in X-Richtung und 0.925 mm bzw. 129 Pixel in Y-Richtung. Dieser Lagefehler kann nun für die Abschätzung der ersten Bedingung, den Maximalabstand korrespondierender Liniensegmente, herangezogen werden. Um ein Suchfenster zu generieren wird der Mittelpunkt des zu matchenden Liniensegments vom Bildkoordinatensystem ins übergeordnete Koordinatensystem transformiert, dabei kann der vorhin abgeschätzte Lagefehler auftreten. Anschließend wird dieser Punkt zurück in das Bildkoordinatensystem des jeweiligen anderen Luftbildes, in dem gerade nach Korrespondenzen gesucht wird, projiziert. Ist hier der Aufnahmewinkel genauso groß und entgegengesetzt, so verdoppelt sich der Fehler im Luftbild und entspricht in diesem schlechtesten Fall dem doppelten Lagefehler. Damit ergibt sich das Ausmaß des Suchfensters, in dem nach Korrespondenzen gesucht wird, welches die **erste Nebenbedingung** darstellt:

Um den vom ersten ins zweite Kamerakoordinatensystem transformierten Liniensegment-Mittelpunkt wird ein rechteckiges Suchfenster definiert, dessen Ausmaß in X- und Y-Richtung jeweils dem zweifachen Lagefehler von 394 bzw. 257 Pixel entspricht. Die Suche nach korrespondierenden Liniensegmenten in diesem Luftbild wird nun auf jene beschränkt, deren Mittelpunkt innerhalb dieses Suchfensters liegt (siehe Abbildung 6.4).

Abbildung 6.4: Beispiel einer erfolgreichen Zuordnung korrespondierender Liniensegmente (rote Linie), wobei die Suche nach Korrespondenzen auf ein Suchfenster (grünes Rechteck) beschränkt wurde.

Nachfolgend wird ein Beispiel gezeigt, bei dem der beschriebene Lagefehler bei der Transformation eines Liniensegment-Mittelpunktes von einem Luftbild in ein anderes auftritt. Wie man anhand von Abbildung 6.5 sieht, schießt der Sehstrahl an der Kante vorbei und trifft das Oberflächenmodell erst auf Straßenniveau, wodurch ein Lagefehler entsteht.

Abbildung 6.5: Darstellungen einer 3D-Ansicht mit Oberflächenmodell bzw. vorläufigem Orthophoto, Luftbild und einem Sehstrahl (hellblaue Linie). Links: Der Sehstrahl geht durch das Projektionszentrum des Luftbilds (roter Punkt) und den Mittelpunkt des im Luftbild detektierten Liniensegments (grüne Linie). Rechts: Bei der Projektion auf das Oberflächenmodell schießt der Sehstrahl an der 3D-Kante (rote Linie) vorbei und trifft die Oberfläche erst auf Straßenniveau.

Mithilfe der Gebäudehöhe von etwa 30 m, die im Oberflächenmodell ersichtlich ist, der Kamerakonstante von 100.5 mm und der Entfernung vom Nadirpunkt (im Luftbild 1 von 1917 Pixel in X-Richtung und 4365 Pixel in Y-Richtung), kann der Lagefehler von 52 Pixel in X-Richtung und 119 Pixel in Y-Richtung berechnet werden, der bei der Transformation ins übergeordnete Koordinatensystem entsteht und in Abbildung 6.5 ersichtlich ist. Bei der Rückprojektion in Luftbild 2 kommt aufgrund der Entfernung vom Nadirpunkt (4099 Pixel in X-Richtung und 2339 Pixel in Y-Richtung) ein Lagefehler von 112 Pixel in X-Richtung und 64 Pixel in Y-Richtung hinzu. In Summe ergibt das einen Lagefehler von 164 Pixel in X-Richtung und 183 Pixel in Y-Richtung (siehe Abbildung 6.6).

Abbildung 6.6: Links: Detektiertes Liniensegment (gelbe Linie) und sein Mittelpunkt (grüner Kreis) in Luftbild 1. Rechts: Von Luftbild 1 in Luftbild 2 transformierter Punkt (roter Kreis), tatsächliche Position des Punktes (grüner Kreis) und Lagefehler (rote Linie) aufgrund der abgerundeten Kante im DOM.

Die berechneten Werte stimmen mit den beobachteten Werten überein, eine Bestätigung dafür, dass die Abschätzung des maximalen Lagefehlers und die darauf basierende Einführung eines Suchfensters sinnvoll ist.

6.2.2.2 Nebenbedingung 2 − kein schleifender Schnitt

Als zweite Nebenbedingung wurde ein Kriterium eingeführt, dass schleifende Schnitte beim Vorwärtsschnitt korrespondierender Liniensegmente verhindert, da diese aus numerischen Gründen Ungenauigkeiten zur Folge hätten, die sich direkt auf die resultierende 3D-Schnittgerade auswirken würden (siehe Abbildung 6.8). Auch hier soll ein Beispiel zur Veranschaulichung der Problematik sowie zur Wahl geeigneter Werte dienen:

Bei einem angenommenen Punktfehler der Endpunkte des Liniensegments im Luftbild von $3 \cdot \sigma_0 = 0.05109$ Pixel (siehe Kapitel 6.3.2 bzw. Formel (6.14)) − ein Fehlerwert, der bei der angenommenen Normalverteilung bei 99.73 % aller Beobachtungen geringer ist − sowie dem daraus resultierenden Winkelfehler der Ebene, die über das Liniensegment und das Projektionszentrum des jeweiligen Luftbildes definiert ist, von 0.00015° (ε in Abbildung 6.7) und einer Flughöhe von 1100 m (H in Abbildung 6.7), ergibt sich bei einem Schnitt mit einer zweiten Ebene, die mit der ersten einen Winkel von 5° | 10° | 15° (ω in Abbildung 6.7) einschließt, ein Fehler in der Höhe der resultierenden Schnittgeraden von 6,7 cm | 3,3 cm | 2,2 cm (ΔH1 bzw. ΔH2 in Abbildung 6.7).

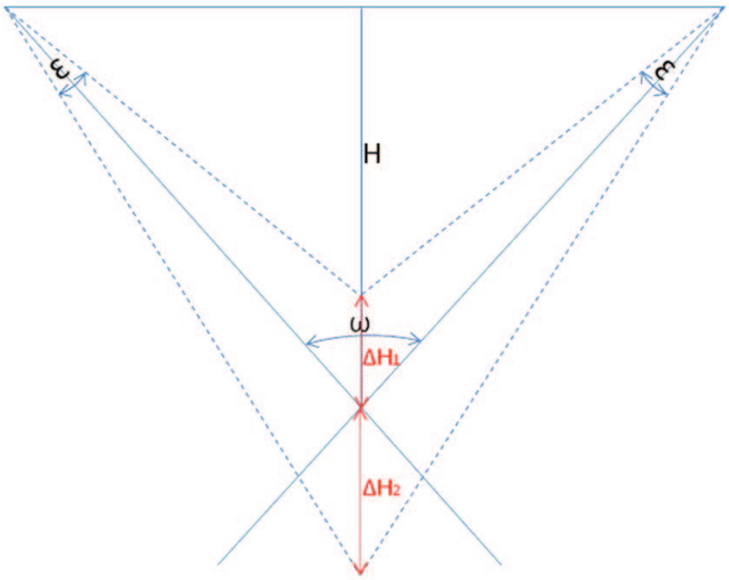

Abbildung 6.7: Schematische Darstellung eines Schnitts mit Winkelfehler der Ebene ε, Flughöhe H, dem Winkel ω, den die Ebenen einschließen und dem daraus resultierenden Fehler in der Höhe ΔH1 bzw. ΔH2, wobei näherungsweise gilt: ΔH1~ΔH2.

Daher wurde, um schleifende Schnitte, bei denen der resultierende Fehler problematisch wäre, als **zweite Nebenbedingung** das Kriterium eingeführt, dass der Winkel zwischen zwei Ebenen mindestens 15° betragen muss, damit die jeweiligen Liniensegmente als korrespondierend angenommen und in der Folge für die Rekonstruktion einer 3D-Kante verwendet werden können.

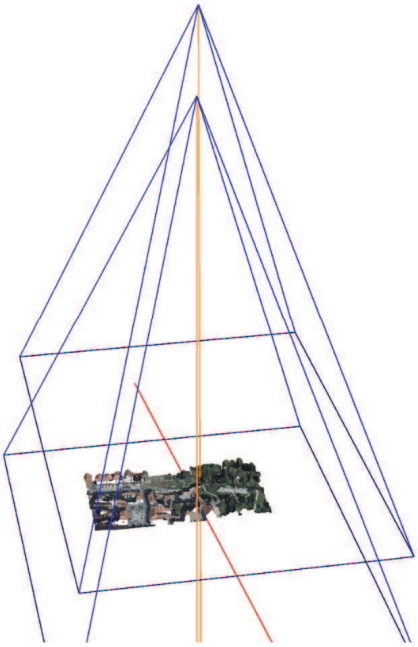

Abbildung 6.8: Beispiel eines schleifenden Schnittes. Die beiden Ebenen, die jeweils über das Projektionszentrum (Spitze der blauen Pyramide) und das Liniensegment im Luftbild definiert sind, schneiden sich in einem sehr spitzen Winkel, wodurch die rekonstruierte Schnittgerade (rote Linie) sehr ungenau ist.

6.2.2.3 Abschließende Bemerkungen zum Matching der Liniensegmente

Es wurde überlegt, als weitere Nebenbedingung einen maximalen Winkel zwischen dem ins andere Kamerakoordinatensystem transformierten Liniensegment 1 und dem Liniensegment 2 einzuführen. Da aber auch Fälle auftreten, bei denen ein Endpunkt des Liniensegments einen hohen Lagefehler aufweist, weil beispielsweise der Sehstrahl an der Kante vorbeischießt (wie bei dem Beispiel in Kapitel 6.2.2.1), der andere Endpunkt aber korrekt transformiert wird (siehe Abbildung 6.9), lässt sich diese Bedingung nicht einführen bzw. wäre der zu wählende Schwellwert für den maximalen Winkel so hoch, dass die Suche nach Korrespondenzen dadurch kaum eingeschränkt wäre.

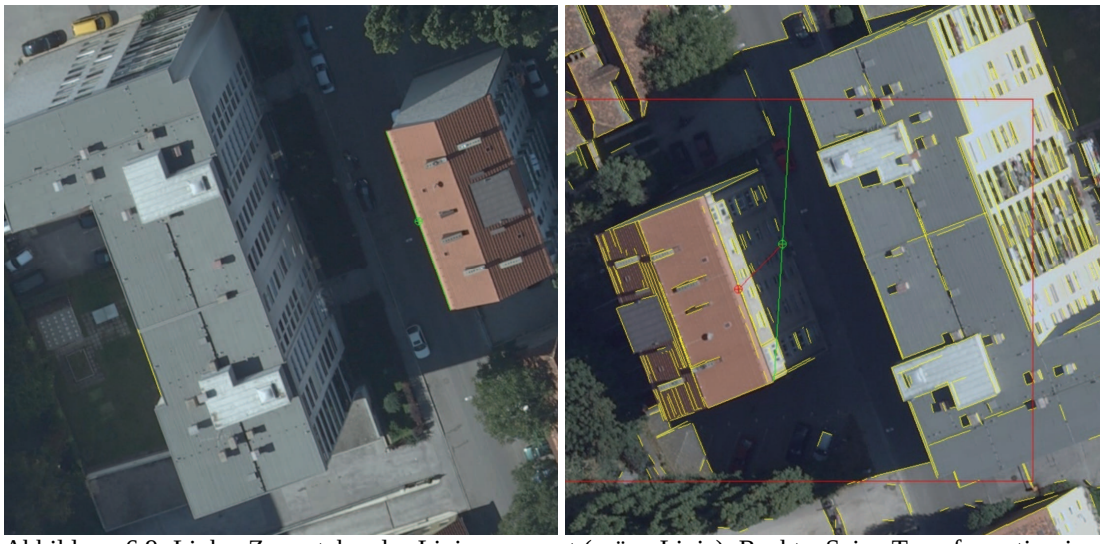

Abbildung 6.9: Links: Zu matchendes Liniensegment (grüne Linie). Rechts: Seine Transformation in ein anderes Luftbild (grüne Linie), bei der ein Endpunkt korrekt transformiert wird, der andere aber einen hohen Lagefehler aufweist und dadurch keine Parallelität zu dem korrespondierenden Liniensegment gegeben ist.

Durch die beiden eingeführten Nebenbedingungen wird die Suche nach korrespondierenden Liniensegmenten auf einige Matching-Kandidaten eingeschränkt. Letztendlich wird dann jener Kandidat als korrespondierend angenommen, der dem anderen Liniensegment im spektralen Merkmalsraum am Nähesten ist, sofern er eine gewisse Minimaldistanz unterschreitet (siehe Kapitel 6.2.1). Durch die Einführung der Nebenbedingungen können einerseits falsche Zuordnungen vermieden werden, andererseits wird aber auch die Effizienz des Algorithmus erheblich gesteigert.

6.3 Rekonstruktion von 3D-Kanten

Die Rekonstruktion von 3D-Kanten erfolgt durch Vorwärtsschnitt korrespondierender Liniensegmente. Dies setzt voraus, dass die innere und äußere Orientierung der jeweiligen Luftbilder gegeben ist. Für jedes Luftbild wird eine Ebene durch das Projektionszentrum und das jeweilige Liniensegment (definiert über Start- und Endpunkt im Bildkoordinatensystem) gelegt. Bei zwei korrespondierenden Liniensegmenten erhält man also zwei Ebenen und kann eindeutig eine Schnittgerade bestimmen. Wurden mehr als zwei korrespondierende Liniensegmente gefunden, ergeben sich entsprechend mehr Ebenen und die Schnittgerade kann nicht mehr eindeutig bestimmt werden. Um dieses überbestimmte Problem zu lösen und die plausibelsten Parameter für die Schnittgerade zu finden, ist eine Ausgleichung notwendig.

6.3.1 Ausgleichung des Vorwärtsschnittes

Die Beobachtungen der Ausgleichung sind die Bildkoordinaten von Start- und Endpunkt von jedem der n Liniensegmente. Da der Algorithmus zur Detektion der Liniensegmente keine Genauigkeitsangaben zu den gefundenen Bildkoordinaten liefert, werden alle Beobachtungen gleich gewichtet. Da auch zu den Parametern der äußeren und inneren Orientierung der Bilder keine Genauigkeitsangaben vorliegen, werden diese nicht als Beobachtungen behandelt, sondern als fest vorgegebene Größen angenommen. Somit ergeben sich pro Ebene bzw. Bild mit Liniensegment 4 Beobachtungen ξ_{Si}, η_{Si}, ξ_{Ei}, η_{Ei} (jeweils zwei Bildkoordinaten des Start- und Endpunktes).

Die Unbekannten der Ausgleichung sind die Parameter der Schnittgeraden im übergeordneten Koordinatensystem. Die übliche Parametrisierung einer Geraden im 3D-Raum ist:

$$\text{g:} \qquad P + \lambda \cdot r = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} + \lambda \cdot \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} \tag{6.1}$$

Da diese Parameterdarstellung einer Geraden im 3D-Raum 6 Parameter enthält, eine Gerade im 3D-Raum aber nur 4 Freiheitsgrade hat, können 2 dieser 6 Parameter festgehalten werden. Es wird erstens der Richtungsvektor *r* so normiert, dass entweder die Koordinate r_x oder r_y − je nachdem, welcher Betrag größer ist − 1 wird und somit festgehalten wird. Als zweites wird der Punkt *P* der Schnittgeraden durch den Schnitt mit einer Ebene, die parallel zur YZ-Ebene (wenn r_x festgehalten wird) bzw. zur XZ-Ebene (wenn r_y festgehalten wird) ist, bestimmt, und damit die p_x- bzw. p_y-Koordinate des Ortsvektors *P* festgehalten. Durch das Festhalten dieser 2 Parameter hat die Ausgleichung nur noch die 4 restlichen Parameter der Gerade als Unbekannte. Im Folgenden wird beim Aufstellen der Formeln davon ausgegangen, dass r_x und p_x festgehalten werden und die Unbekannten somit r_y, r_z, p_y und p_z sind.

Um sowohl Richtung als auch Position simultan auszugleichen, werden zwei Bedingungsgleichungen eingeführt, welche den funktionalen Zusammenhang zwischen Beobachtungen und Unbekannten darstellen.

1. Bedingungsgleichung:

$$n_i^T \cdot r = 0 \qquad \text{bzw.} \quad n_{xi} \cdot r_x + n_{yi} \cdot r_y + n_{zi} \cdot r_z = 0 \tag{6.2}$$

Der Normalvektor n_i jeder Ebene soll zum Richtungsvektor *r* der Schnittgeraden orthogonal stehen und daher muss ihr Skalarprodukt 0 sein. Diese Bedingungsgleichung enthält die beiden Unbekannten des Richtungsvektors der Schnittgeraden r_y und r_z, denn r_x ist wie zuvor beschrieben 1.

2. Bedingungsgleichung:

$$n_i^T \cdot P - d_i = 0 \qquad \text{bzw.} \quad n_{xi} \cdot p_x + n_{yi} \cdot p_y + n_{zi} \cdot p_z - d_i = 0 \tag{6.3}$$

Hierbei bedient man sich der Hesseschen Normalform, mit der sich eine Ebene im dreidimensionalen Raum definieren, beziehungsweise der Abstand eines beliebigen Punktes zu dieser Ebene berechnen lässt. Das Skalarprodukt des Normalvektors n_i einer Ebene und des Ortsvektors eines Punktes *P* dieser Ebene ergibt den Parameter d_i derselben Ebene, welcher den Abstand der Ebene zum Ursprung repräsentiert ($d_i / |n_i| =$ Abstand zum Ursprung). Die Ebene besteht aus denjenigen Punkten, die diese Gleichung erfüllen. Da die Schnittgerade in jeder der zu schneidenden Ebenen liegen soll, muss obige Bedingung nicht nur für einen Punkt der jeweiligen Ebene sondern auch für einen Punkt *P* der Schnittgeraden gelten. Diese Bedingungsgleichung enthält also die beiden Unbekannten des Ortsvektors *P* eines Punktes der Schnittgeraden p_y und p_z, denn p_x wird wie zuvor beschrieben, durch den Schnitt mit einer YZ-Ebene bestimmt.

Die Ausgleichung umfasst also 4 Unbekannte sowie pro Ebene 4 Beobachtungen und 2 Bedingungsgleichungen, die das funktionale Modell bilden, sprich den funktionalen Zusammenhang von Unbekannten und Beobachtungen darstellen. Da die Funktionen nicht jeweils nur eine einzige Beobachtung beinhalten (in den Normalvektoren stehen jeweils alle 4 Beobachtungen der Ebene), liegt keine Ausgleichung nach vermittelnden Beobachtungen vor. Es gibt ebenfalls keine Bedingungen zwischen den Unbekannten. Somit liegt der Allgemeinfall der Ausgleichungsrechnung vor, die Ausgleichung nach bedingten Beobachtungen mit Verbesserungsgleichungen, welcher sich mit dem Gauß-Helmert-Modell der Ausgleichung lösen lässt [Navratil, 2012].

Durch das Ableiten der $2 \cdot \mathrm{n}$ Bedingungsgleichungen nach den 4 Unbekannten ergibt sich die A-Matrix:

$$A = \begin{bmatrix} n_{y1} & n_{z1} & 0 & 0 \\ 0 & 0 & n_{y1} & n_{z1} \\ & & \vdots & \\ & & \vdots & \\ n_{yn} & n_{zn} & 0 & 0 \\ 0 & 0 & n_{yn} & n_{zn} \end{bmatrix} \tag{6.4}$$

Durch das Ableiten der $2 \cdot n$ Bedingungsgleichungen nach den $4 \cdot n$ Beobachtungen ergibt sich die B-Matrix als Blockmatrix, die aus mehreren 2 mal 4 Blöcken besteht, wobei die Blöcke außerhalb der Diagonalen Nullmatrizen sind:

$$B = \begin{bmatrix} \begin{array}{|c|c|c|c|} \hline (R0_i \times OE_i) \cdot r & (R1_i \times OE_i) \cdot r & (OS_i \times R0_i) \cdot r & (OS_i \times R1_i) \cdot r \\ \hline (R0_i \times OE_i) \cdot P - (R0_i \times OE_i) \cdot O_i & (R1_i \times OE_i) \cdot P - (R1_i \times OE_i) \cdot O_i & (OS_i \times R0_i) \cdot P - (OS_i \times R0_i) \cdot O_i & (OS_i \times R1_i) \cdot P - (OS_i \times R1_i) \cdot O_i \\ \hline \end{array} & \cdots & \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \\ \vdots & \ddots & \\ \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} & & \ddots \end{bmatrix}$$

(6.5)

wobei gilt:

$$\begin{aligned} n_i &= OS_i \times OE_i = [S_i - O_i] \times [E_i - O_i] = \\ &= \left[\left(R_i^T \begin{pmatrix} \xi_{Si} \\ \eta_{Si} \\ f \end{pmatrix} + O_i \right) - O_i \right] \times \left[\left(R_i^T \begin{pmatrix} \xi_{Ei} \\ \eta_{Ei} \\ f \end{pmatrix} + O_i \right) - O_i \right] \\ &= \left[R_i^T \begin{pmatrix} \xi_{Si} \\ \eta_{Si} \\ f \end{pmatrix} \right] \times \left[R_i^T \begin{pmatrix} \xi_{Ei} \\ \eta_{Ei} \\ f \end{pmatrix} \right] \end{aligned} \tag{6.6}$$

$$d_i = n_i \cdot O_i \tag{6.7}$$

mit
- O_i … Projektionszentrum im übergeordneten Koordinatensystem
- S_i … Startpunkt im übergeordneten Koordinatensystem
- E_i … Endpunkt im übergeordneten Koordinatensystem
- R_i^T … 3 mal 3 Rotationsmatrix zur Transformation vom Bildkoordinatensystem ins übergeordnete Koordinatensystem
- $R0_i, R1_i, R2_i$ … Spalten der Rotationsmatrix R_i^T
- ξ_{Si}, η_{Si} … Bildkoordinaten vom Startpunkt
- ξ_{Ei}, η_{Ei} … Bildkoordinaten vom Endpunkt
- f … Kamerakonstante
- r … Richtungsvektor der Schnittgeraden
- P … Punkt auf der Schnittgeraden

Um eine Näherung der Schnittgeraden bzw. der Unbekannten X_0 zu bestimmen, werden jene 2 Ebenen miteinander geschnitten, die sich am ehesten im rechten Winkel schneiden. Über den funktionalen Zusammenhang lässt sich dann der Widerspruchsvektor w bestimmen, wobei für jede Ebene die entsprechenden Beobachtungen und die genäherten Unbekannten in die beiden Bedingungsleichungen eingesetzt werden:

$$w = \varphi(L, X_0) \tag{6.8}$$

Da keine Genauigkeitsangaben zu den Beobachtungen und keine stochastischen Abhängigkeiten zwischen diesen vorliegen, vereinfachen sich die Kovarianzmatrix Σ_{LL}, die Kofaktormatrix Q_{LL} und die Gewichtsmatrix P zur Einheitsmatrix I.

Nun lässt sich die Extremwertaufgabe $v^T Pv \rightarrow Min$, also das Minimieren der Verbesserungsquadrate, mit der Nebenbedingung $Ax+Bv+w=0$ folgendermaßen lösen:

$$\begin{pmatrix} k \\ x \end{pmatrix} = \begin{pmatrix} BP^{-1}B^T & A \\ A^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} -w \\ 0 \end{pmatrix} \tag{6.9}$$

$$v = P^{-1}B^T k \tag{6.10}$$

Über den gekürzten Parametervektor x und die Verbesserungen v erhält man letztendlich die ausgeglichenen Unbekannten $\hat{X}$ und die ausgeglichenen Beobachtungen $\hat{L}$:

$$\hat{X} = X_0 + x \tag{6.11}$$

$$\hat{L} = L + v \tag{6.12}$$

Als Ergebnis der Ausgleichung erhält man unter der Annahme, dass die vorliegenden Daten normalverteilt, also nicht mit systematischen oder groben Fehlern behaftet sind, eine statistisch gesehen optimale Schnittgerade, da die Quadrate der Fehler in den direkt gemessenen Beobachtungen (den Bildkoordinaten der Start- und Endpunkte der Liniensegmente) minimiert wurden.

6.3.2 Erkennen und Eliminieren von Ausreißern

Da die Ausgleichung nur dann eine optimale Schnittgerade als Ergebnis liefern kann, wenn keine groben Fehler in den Beobachtungen vorliegen, ist es sinnvoll Tests durchzuführen, um mögliche grobe Fehler aufdecken und eliminieren zu können. Ein solcher grober Fehler liegt beispielsweise vor, wenn beim Matching Liniensegmente aus unterschiedlichen Bildern zugeordnet werden, obwohl sie nicht die gleiche Kante abbilden. Es ist mit dem Vorkommen solcher groben Fehler zu rechnen, da trotz der komplexen Heuristik, wie sie in Kapitel 6.2 beschrieben ist, bei mehreren tausenden Kanten, die gematcht werden, höchstwahrscheinlich auch falsche zugeordnete Korrespondenzen auftreten.

Als Standardgröße zur Detektion von groben Fehlern haben sich die normierten Verbesserungen nv_i etabliert:

$$nv_i = {v_i}/{\sigma_0} \tag{6.13}$$

mit v_i … Verbesserung der jeweiligen Beobachtung

σ_0 … Standardabweichung der Verbesserungen
bzw. mittlere Punktgenauigkeit der Endpunkte der Liniensegmente

Das σ_0 als Standardabweichung der Verbesserungen entspricht der mittleren Punktgenauigkeit der Endpunkte der detektierten Liniensegmente und wurde a posteriori mithilfe von Stichproben abgeschätzt. Dabei wurden 10 Matches, die jeweils mindestens je 3 Liniensegmente enthalten, zufällig ausgewählt. Genau genommen wurden dabei Matches mit 2 mal 3, 5 mal 4, 2 mal 6 und 1 mal 7 Korrespondenzen beziehungsweise in Summe 45 Liniensegmente oder 180 Beobachtungen herangezogen. Sie wurden optisch kontrolliert, um sicherzugehen, dass sie keine groben Fehler aufweisen und dann einer Ausgleichung (wie in Kapitel 6.3.1 beschrieben) unterzogen. Anschließend wurden die Verbesserungen statistisch analysiert und σ_0 bestimmt.

Das arithmetische Mittel der Verbesserungen sollte gegen 0 gehen, da keine systematischen Fehlereinflüsse bekannt sind und der Erwartungswert μ daher 0 beträgt. Tatsächlich beträgt es -0.00026 Pixel, was bei einem Stichprobenumfang von 180 Beobachtungen in der zu erwartenden Schwankungsbreite liegt.

Aufgrund des bekannten Erwartungswerts μ von 0, lässt sich die Standardabweichung der Verbesserungen σ_0 über das quadratische Mittel berechnen:

$$\sigma_0 = \sqrt{\sum (v_i^2 - \mu)/n} = \sqrt{\sum v_i^2/n} = 0.01703 \text{ Pixel} \tag{6.14}$$

Mit dem bekannten σ_0 können nun bei jeder Ausgleichung die normierten Verbesserungen nv_i durch Division der Verbesserungen v_i mit σ_0 (Formel (6.13)) bestimmt werden. Somit erhält man eine standardisierte Normalverteilung der Verbesserungen, bei der gilt: Erwartungswert $\mu = 0$ und Standardabweichung $\sigma = 1$.

Bei einer standardisierten Normalverteilung sind 99.9937% aller Daten im Wertebereich von $\mu \pm 4 \cdot \sigma$. Dies wurde auch als Kriterium für Ausreißer gewählt, sprich ein Liniensegment wurde dann aus der Ausgleichung eliminiert, wenn eine seiner vier Beobachtungen eine Verbesserung aufweist, die nicht in diesem Wertebereich liegt. Die Ausgleichung wird dann mit den übrigen Liniensegmenten wiederholt, wobei allerdings immer nur ein Liniensegment eliminiert wird und zwar jenes mit der Verbesserung, die am weitesten von diesem Wertebereich abweicht.

Wenn keine Ausreißer mehr vorliegen, sprich alle Verbesserungen innerhalb des vorgegebenen Wertebereichs liegen, wird zur Kontrolle der Ausgleichung die Hauptprobe gemacht. Dabei werden die ausgeglichenen Beobachtungen und Unbekannten in die Bedingungsgleichungen eingesetzt, um zu überprüfen, ob der funktionale Zusammenhang genau genug erfüllt ist. Geht die Hauptprobe nicht auf, wird die rekonstruierte 3D-Kante eliminiert, da man davon ausgehen muss, dass sie fehlerhaft ist.

$$\varphi(\hat{L}, \hat{X}) \cong 0 \tag{6.15}$$

Zur Veranschaulichung wird ein Beispiel einer erfolgreichen Eliminierung von Ausreißern gezeigt (siehe Abbildung 6.10 und Abbildung 6.11).

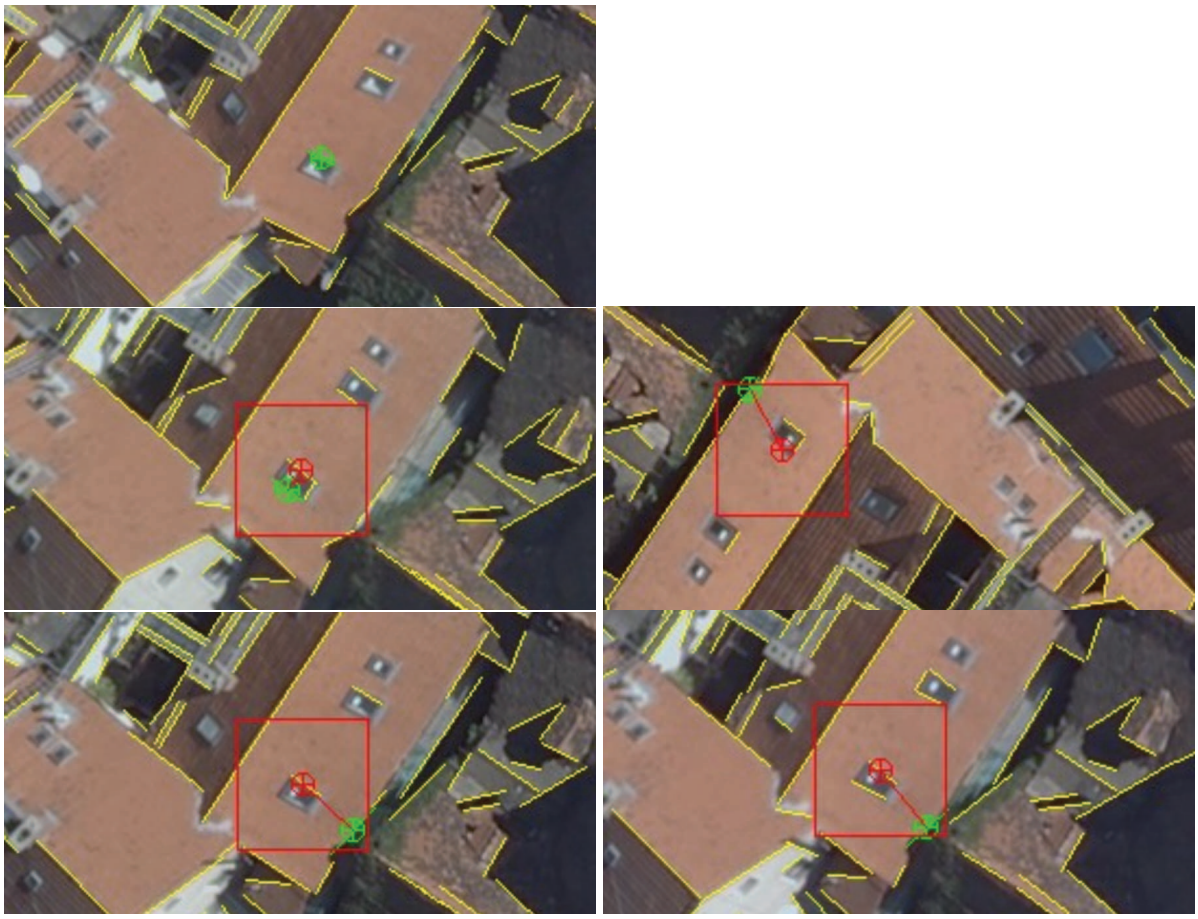

Abbildung 6.10: Matchingergebnis ohne Eliminierung der Ausreißer: gematchtes Liniensegment (grüne Linie), welches die Kante eines Dachfensters abbildet (oben) und 4 gefundene Korrespondenzen (grüne Linie), von denen nur die erste korrekt ist (Mitte links). Die übrigen bilden die Traufenkante ab (Mitte rechts und unten).

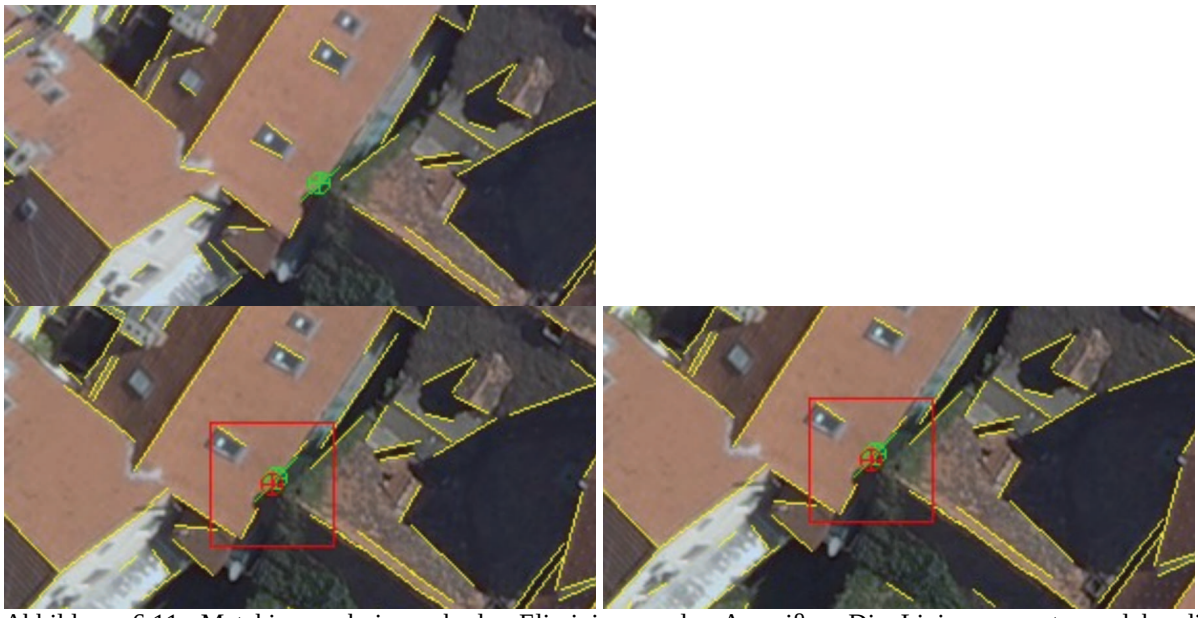

Abbildung 6.11: Matchingergebnis nach der Eliminierung der Ausreißer: Die Liniensegmente, welche die Dachfensterkante abbilden wurden als Ausreißer erkannt und eliminiert. Das nun gematchte Liniensegment (oben) bildet, so wie auch die 2 gefundenen korrekten Korrespondenzen (unten), die Traufenkante ab.

6.3.3 Bestimmung von 3D-Kanten

Nachdem im Zuge einer Ausgleichung eine 3D-Schnittgerade bestimmt wurde, geht es nun darum, auf dieser eine endliche 3D-Kante, die vom Ausmaß her den gematchten Liniensegmenten entspricht, zu definieren. Da bei der Detektion der Liniensegmente Kanten oft nur lückenhaft, teils in mehreren unvollständigen Segmenten abgebildet werden, wurde der Ansatz gewählt, dass von den gematchten Liniensegmenten das längste auf die Schnittgerade projiziert und damit das Ausmaß der 3D-Kante bestimmt wird (siehe Abbildung 6.12). Da die Luftbilder im Zuge einer Befliegung entstanden sind und daher einen ähnlichen Maßstab aufweisen, ist davon auszugehen, dass das längste Liniensegment den größten Teil der Kante abbildet.

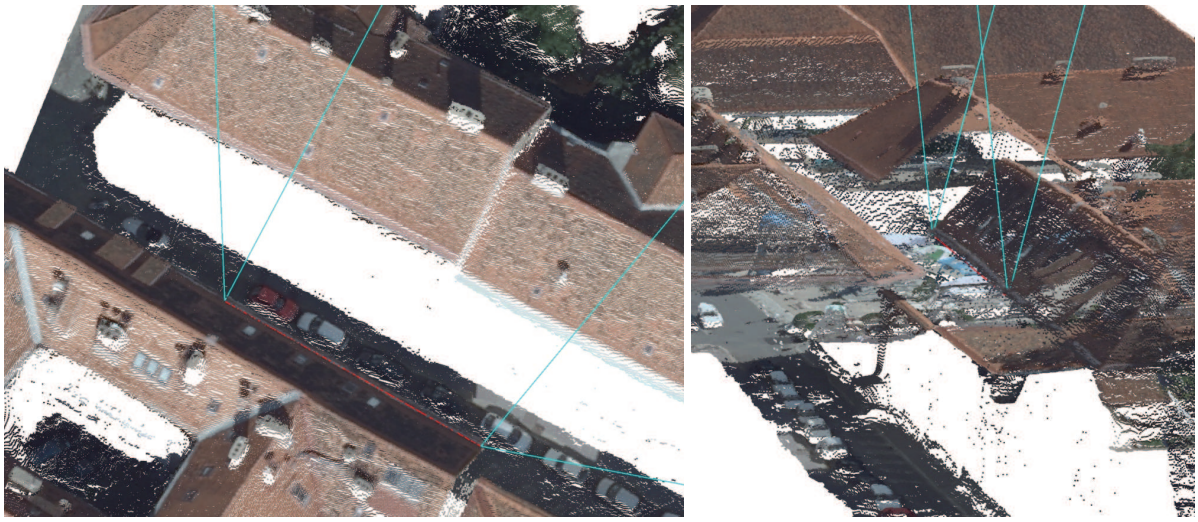

Abbildung 6.12: Rekonstruktion einer 3D-Traufenkante (rote Linie), eingezeichnet in eine Visualisierung der Orthophoto-Punktwolke. Die blauen Linien stellen Strahlen zweier Luftbilder dar, die vom Projektionszentrum durch die Endpunkte der entsprechenden Liniensegmente gehen.

7 Ergebnisse

7.1 Ergebnisse mittels der entwickelten Implementierung

Es werden nun die Ergebnisse gezeigt, die der entwickelte Algorithmus zum Matching der Liniensegmente und zur anschließenden Rekonstruktion von 3D-Kanten bei den gegebenen Daten, also den Luftbildern, sowie einem groben Höhenmodell für Graz, erzielt. Im Folgenden werden für die Ergebnisse charakteristische Beispiele gezeigt und zwar vor allem jene, bei denen das Matching falsche Liniensegmentkorrespondenzen bestimmt hat. Dabei sollen mögliche Ursachen für die falschen Zuordnungen angegeben und damit auch die Grenzen des entwickelten Matching-Algorithmus aufgezeigt werden.

7.1.1 Symbolik der Abbildungen

Bei den vorliegenden Abbildungen, welche die Matching-Ergebnisse zeigen, wurde eine einheitliche Symbolik verwendet. Die zu matchenden bzw. gematchten Liniensegmente sind als grüne Linie dargestellt, ihr Mittelpunkt als grüner Kreis. Der rote Kreis zeigt den transformierten Mittelpunkt des zu matchenden Liniensegments und das rote Rechteck symbolisiert das Suchfenster, in welchem nach Korrespondenzen gesucht wird. Die rote Linie stellt den Abstand zwischen dem roten und dem grünen Kreis dar.

7.1.2 Beispiel 1 – ähnliche sich wiederholende Strukturen

Im vorliegenden Beispiel (Abbildung 7.1) liegt eine Szene vor, bei der sich sehr ähnliche Strukturen mehrmals wiederholen. In diesem Fall sind es benachbarte Firste und Kehle, die sich bezüglich ihrer spektralen Charakteristik kaum voneinander unterscheiden und für den Deskriptor daher nicht differenzierbar sind. Im Allgemeinen stoßen viele Matching-Algorithmen bei derartigen Szenen an ihre Grenzen. Mit Deskriptoren, die nur auf spektralen Informationen beruhen, lassen sich die Liniensegmente aufgrund ihrer sehr hohen Ähnlichkeit de facto nicht unterscheiden. Geometrische Information wie die Position lassen sich im vorliegenden Fall über die Transformation des Liniensegments in die anderen Luftbilder mithilfe des groben Höhenmodells abschätzen. Diese wäre hier auch relativ genau (siehe roter Kreis in Abbildung 7.1) und könnte in diesem Fall verwendet werden, um die korrekten Korrespondenzen zu finden. Im Allgemeinen ist diese Information aber nicht zuverlässig, da sie auf dem unzuverlässigen, groben Höhenmodell beruht und es können Fälle auftreten, bei denen die Transformation sehr ungenau ist (ein Beispiel dafür ist in Kapitel 6.2.2.1 gegeben). Außerdem ist es ja schließlich das Ziel, dieses Höhenmodell an genau jenen Kanten zu verbessern, wo noch Ungenauigkeiten vorliegen. Daher ist es notwendig, die über die Transformation abgeschätzte Position als ungenau zu betrachten und ein größeres Suchfenster (eine sinnvolle Größe wurde in Kapitel 6.2.2.2 bestimmt) zu wählen. Andere geometrische Informationen wie die Orientierung der Liniensegmente lassen sich auch nicht zuverlässig bestimmen (siehe Kapitel 6.2.2.3) und würden im vorliegenden Fall auch gar keine Abhilfe schaffen, ebenso wie die Länge der Liniensegmente. Vielleicht lässt sich über die Einführung von Epipolarbedingungen, welche auf der Orientierung der Luftbilder basieren, die Unterscheidbarkeit von Liniensegmenten derartiger Szenen ermöglichen. Der in Kapitel 4.1.3

beschriebene Algorithmus beruht auf diesem Ansatz, die Ergebnisse werden in Kapitel 7.2 vorgestellt.

Abbildung 7.1: Liniensegment (links, grüne Linie) und seine falsch bestimmten Korrespondenzen (übrige Bilder, grüne Linie) in drei anderen Luftbildern. Der rote Kreis zeigt den transformierten Mittelpunkt des zu matchenden Liniensegments, der grüne Kreis den Mittelpunkt des gematchten Liniensegments und das rote Rechteck symbolisiert das Suchfenster indem nach Korrespondenzen gesucht wurde.

7.1.3 Beispiel 2 – Richtung der Liniensegmente

Dieses Beispiel offenbart eine weitere Schwäche des Algorithmus. Um jeweils die gleichen Farbprofile vergleichen zu können, werden die Liniensegmente so gerichtet, dass das dunklere Profil jeweils rechts ist. In den allermeisten Fällen ist dies in allen Luftbildern auf derselben Seite und somit werden die gleichen Farbprofile verglichen. Es können Ausnahmefälle auftreten, wo dies nicht gewährleistet ist und dadurch das Matching nicht erfolgreich sein kann. Im vorliegenden Beispiel, so wie auch in allen anderen gezeigten, befindet sich das dunklere Profil in allen Luftbildern aber jeweils auf der gleichen Seite. Dafür scheitert der Algorithmus hier an einem etwas anderem Sachverhalt. Und zwar liegt eine Fläche mit einem relativ konstanten Farbverlauf vor, die von einer anderen, welche ebenso eine relativ konstante spektrale Reflexionscharakteristik aufweist, umgeben ist. Im vorliegendem Fall handelt es sich dabei um eine Wiese, von der sich ein kleiner Bereich im Schlagschatten befindet, die umgebende Fläche allerdings nicht. Bei den detektierten Liniensegmenten bildet das rechte Profil daher immer den im Schatten liegenden Bereich ab und das linke die von der Sonne bestrahlte Wiese, und zwar auch dann, wenn es sich um gegenüberliegende Liniensegmente handelt. Somit können die Liniensegmente, welche diese beiden Bereiche abtrennen, mit dem Deskriptor kaum unterschieden werden und es ist nicht verwunderlich, dass falsche Korrespondenzen bestimmt werden (siehe Abbildung 7.2).

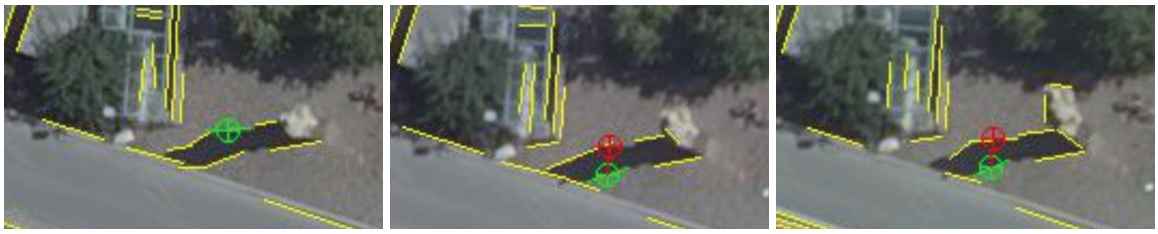

Abbildung 7.2: Liniensegment (links, grüne Linie) und seine falsch bestimmten Korrespondenzen (übrige Bilder, grüne Linie) in zwei anderen Luftbildern.

7.1.4 Beispiel 3 – Farbprofil nur auf einer Seite vergleichbar

Im Folgenden seien Beispiele von Matching-Ergebnissen von Traufenkanten gezeigt, bei denen jeweils nur ein Farbprofil vergleichbar ist, weil im Luftbild mit dem zu matchenden Liniensegment die Fassade nicht sichtar ist, in den übrigen Luftbildern hingegen schon. Im ersten Beispiel wird das Liniensegment auf der Traufenkante im Bild ohne sichtbare Fassade mit Liniensegmenten auf einem Fenster in den Bildern mit sichtbarer Fassade gematcht (siehe Abbildung 7.3). Offensichtlich weist zumindest ein Profil eine ähnliche spektrale Charakteristik auf, dies dürfte jenes sein, dass die Dachrinne bzw. den Schlagschattenbereich des Fensters abbildet und jeweils sehr dunkel ist.

Abbildung 7.3: Liniensegment auf der Traufenkante (links), wird mit Liniensegmenten auf einem Fenster gematcht (Mitte und links).

Als zweites wird ein Beispiel mit demselben Ausschnitt gezeigt, bei dem ebenfalls ein Liniensegment auf der Traufenkante gematcht werden soll. Hier liefert das Matching allerdings ein korrektes Ergebnis, obwohl auch diesmal nur ein Profil vergleichbar ist, denn die dachseitigen Profile weisen eine ähnliche spektrale Charakteristik auf (siehe Abbildung 7.4).

Abbildung 7.4: Korrektes Matching von Liniensegmenten auf der Traufenkante, obwohl nur das dachseitige Profil vergleichbar ist.

7.1.5 Beispiel 4 – Doppelkanten

Doppelkanten können mit dem Matching-Algorithmus im Allgemeinen gut unterschieden werden. Im vorliegenden Beispiel wird die detektierte Traufenkante auch in zwei Luftbildern wiedergefunden (siehe Abbildung 7.5), obwohl aufgrund der Sichtbarkeit der Fassade in diesen Luftbildern nur ein Farbprofil der Kante vergleichbar ist. Dennoch konnte die korrespondierende Kante bestimmt werden, da das vergleichbare Profil eine sehr ähnliche spektrale Charakteristik aufweist und sich die Profile der Doppelkante deutlich in ihren spektralen Eigenschaften unterscheiden.

Abbildung 7.5: Detektiertes Liniensegment der Traufenkante (oben) wird in zwei weiteren Luftbildern wiedergefunden (unten links und unten rechts).

Auch die Kante parallel zu der Traufe wurde korrekt gematcht (siehe Abbildung 7.6), schließlich weist sie in den vorliegenden Luftbildern eine spektrale Charakteristik auf, die sich von denen anderer umliegender Kanten innerhalb des Suchfensters deutlich unterscheidet. Auch sind von dieser Kante beide Farbprofile vergleichbar.

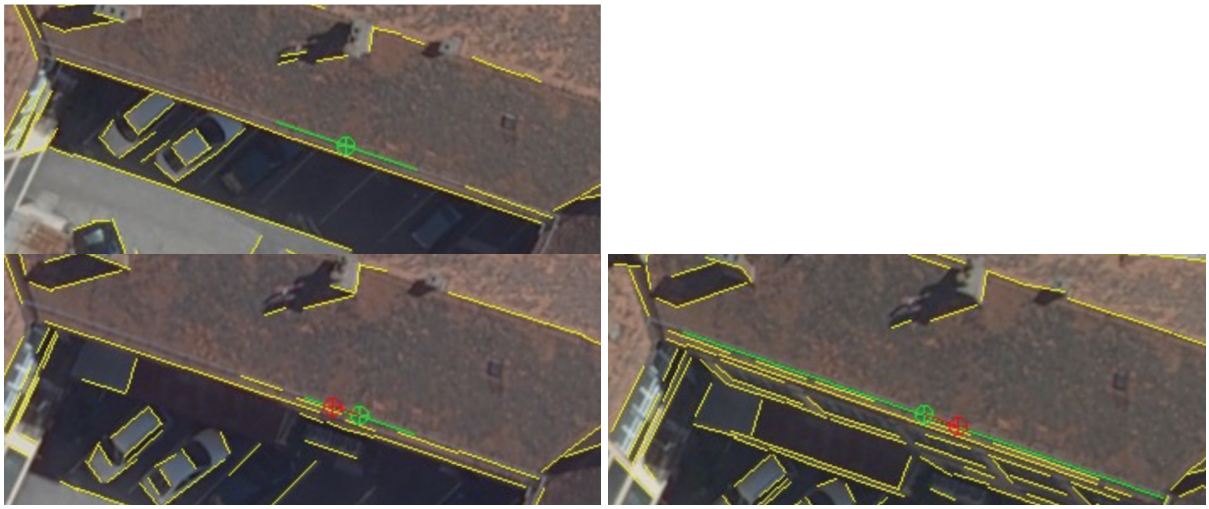

Abbildung 7.6: Detektiertes Liniensegment parallel zur Traufe (oben) wird in zwei weiteren Luftbildern wiedergefunden (unten links und unten rechts).

Dennoch kommt es immer wieder vor, dass Doppelkanten nicht korrekt gematcht werden (siehe Abbildung 7.7). Im nachfolgenden Beispiel fällt im ersten Luftbild auf, dass die Kante parallel zur Traufe in zwei nebeneinander liegenden Liniensegmenten abgebildet wird, obwohl es sich wie erwähnt nur um eine Kante handelt. Das Matching der oberen dieser beiden Liniensegmente schlägt, wie man sieht, fehl. Ein Grund dafür sind die winkelabhängigen Reflexionsunterschiede, den die Luftbilder aufweisen, da sie von einem anderen Flugstreifen stammen. Dies führt dazu, dass nun die spektrale Ähnlichkeit mit den Liniensegmenten, die in den anderen Luftbildern auf der Traufenkante detektiert wurden, am höchsten ist. Denn sowohl die dachseitigen Profile als auch jene die den oberen Bereich der Fassade bzw. die Doppelkante abbilden, sehen sehr ähnlich aus.

Abbildung 7.7: Die Kante parallel zur Traufe wird im ersten Luftbild (oben) auffälligerweise doppelt detektiert. In den anderen Luftbildern (Mitte und unten) wird sie fälschlicherweise mit dem Liniensegment der Traufe gematcht. Ein Grund dafür sind winkelabhängige Reflexionsunterschiede.

7.1.6 Beipsiel 5 – Probleme bei der Kantendetektion

Limitierend für den Matching-Algorithmus können auch Probleme bei der Kantendetektion sein. Einerseits kommt es vor, dass Kanten nicht in allen Luftbildern detektiert werden, andererseits ist es aber auch möglich, dass in den unterschiedlichen Luftbildern zwar dieselben Kanten detektiert werden, sie dabei aber entweder in mehrere Teilsegmente zerstückelt sind oder nicht exakt auf der Kante liegen und daher Unterschiede sowohl in der Position als auch in der Orientierung aufweisen (siehe Abbildung 7.8). Dies hat zur Folge, dass die detektierten Segmente selbst dann, wenn sie beim Matching einander zugeordnet werden, möglicherweise im Zuge der 3D-Rekonstruktion bei der Ausgleichung als Ausreißer detektiert und daher eliminiert werden.

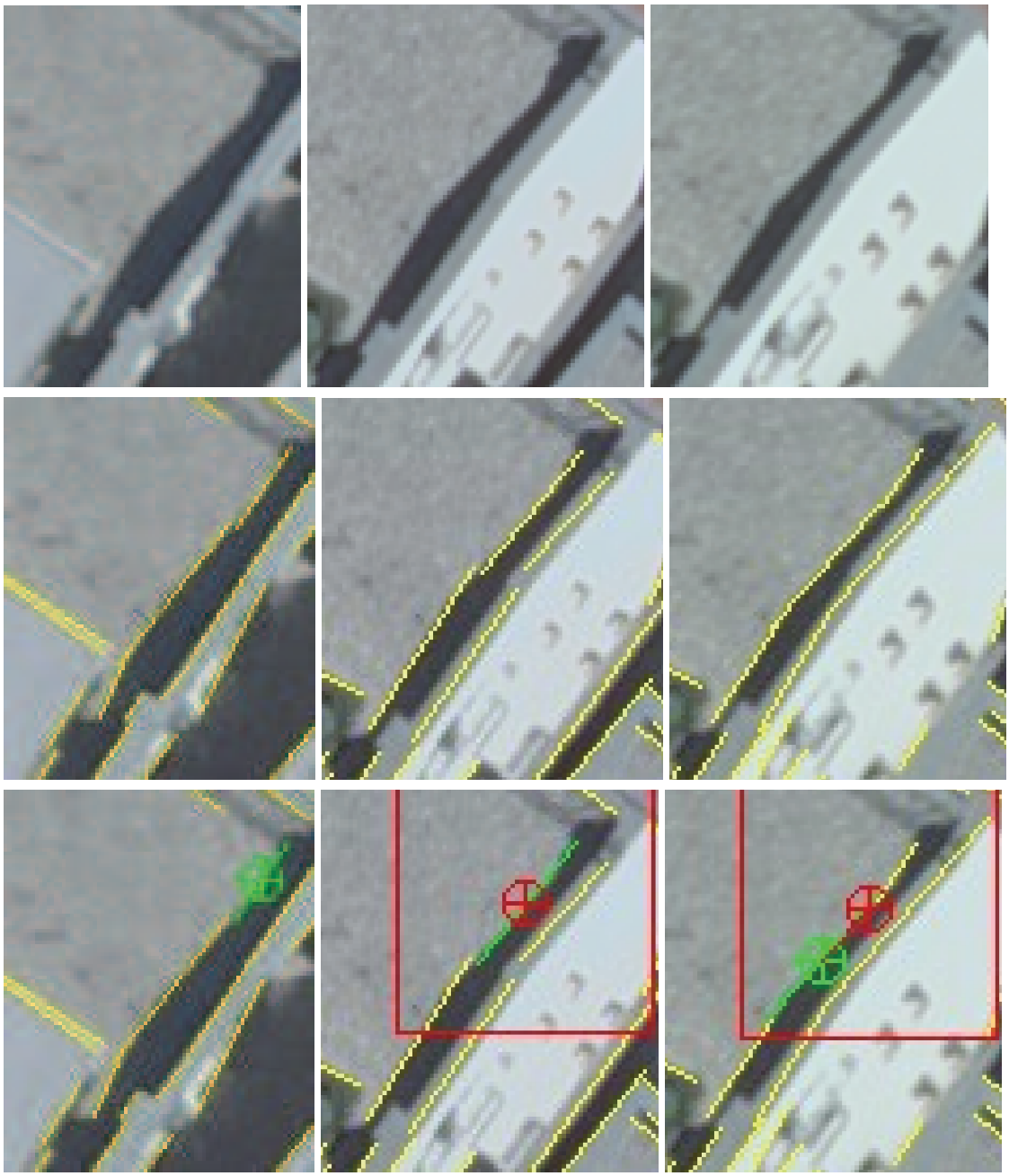

Abbildung 7.8: Beispiel für die unterschiedliche Detektion von Liniensegmenten auf derselben Kante in verschiedenen Luftbildern. Oben: Luftbilder; Mitte: Luftbilder mit detektierten Liniensegmenten; Unten: Luftbilder mit korrespondierenden Liniensegmenten.

7.1.7 Beispiel 6 – erfolgreiches Matching und 3D-Rekonstruktion

Abschließend sei noch ein Beispiel gezeigt, bei dem eine Firstkante erfolgreich gematcht und dreidimensional rekonstruiert wird (siehe Abbildung 7.9). In allen Luftbildern wird die Kante gleichermaßen detektiert, wobei sogar die Länge der Liniensegmente fast ident ist. Die beiden Farbprofile weisen aufgrund der winkelabhängigen Reflexionseigenschaften des Dachs charakteristische spektrale Eigenschaften auf, die in allen Luftbildern sehr ähnlich sind, wodurch die Korrespondenzen problemlos bestimmt werden können. Die anschließende 3D-Rekonstruktion der Firstkante erfolgt im Zuge einer Ausgleichung und liefert ein sehr genaues Ergebnis, welches hier in ein vorläufiges Orthophoto, das mit dem groben DOM berechnet wurde und daher an Kanten − wie dem First − noch größere Ungenauigkeiten beinhalten kann, eingezeichnet ist.

Abbildung 7.9: Matching und 3D-Rekonstruktion einer Firstkante.

7.2 Ergebnisse mittels der Implementierung von Werner und Zisserman (2002)

Beim Testen der MATLAB-Implementierung zum Matching von Liniensegmenten und zur Rekonstruktion von 3D-Kanten (siehe Kapitel 4.1.3) wurden insgesamt 14 Luftbilder, die von den drei benachbarten Flugstreifen A, B und C stammen, verwendet. Von diesen Luftbildern wiederum wurden jeweils nur kleine Kacheln ausgewählt, welche mindestens eine Überschneidung von 90% bezüglich der abgebildeten Szene aufweisen (siehe Abbildung 7.10). Diese Einschränkung war notwendig, da die verwendete MATLAB-Software keine ganzen Luftbilder einlesen und als Matrix speichern kann, da jene zu groß wären. Die Liniensegmente werden in jedem Luftbild mit einem Canny-Kantendetektor bestimmt, anschließend wird ein Matching der Liniensegmente durchgeführt und basierend auf den gefundenen Korrespondenzen 3D-Kanten rekonstruiert und im VRML-Format exportiert. Diese sind in den Abbildungen der gezeigten Ergebnisse dargestellt.

Abbildung 7.10: Jeweils die erste (oben) und letzte (unten) Luftbildkachel von drei benachbarten Flugstreifen A (links), B (Mitte) und C (rechts). Alle bilden in etwa dieselbe Szene ab, der mittlere Streifen ist gegenüber den anderen bedingt durch die Flugrichtung um 180° verdreht.

Die Orientierung der Luftbilder wird als Input benötigt, um sowohl Epipolarbedingungen als auch Rückprojektionsbedingungen für die Liniensegmente aufzustellen. Diese Bedingungen müssen nicht exakt erfüllt sein, sondern können bedingt durch die Ungenauigkeit der gegebenen Orientierungen kleine Abweichungen aufweisen. Für diese wurden verschiedene Schwellwerte gesetzt und die Implementierung damit getestet. Im Allgemeinen muss man aber sagen, dass bei allen gewählten Parametereinstellungen die erzielten Ergebnisse hinsichtlich ihrer Eignung, die abgebildete Szene flächendeckend so zu rekonstruieren, dass damit ein Höhenmodell verbessert werden kann, nicht befriedigend sind. Im Folgenden seien dennoch jene gezeigt, die am ehesten brauchbar sind.

Zunächst wurde die Implementierung mit den Luftbildkacheln eines Flugstreifens und den Standardparametereinstellungen getestet. Dabei wurde der Streifen C gewählt, da die verwendeten Kacheln die größte Überschneidung bezüglich der abgebildeten Szene aufweisen und damit die besten Ergebnisse zu erwarten sind. Das in Abbildung 7.11 dargestellte Ergebnis der rekonstruierten 3D-Kanten lässt die in den Luftbildern abgebildete Szene erkennen. Vor allem der Straßenzug mit der Kreuzung sowie nördlich davon die großen

Häuser mit Innenhöfen sind klar ersichtlich. Die kleineren Häuser südlich der Straße sind hingegen nicht so gut rekonstruiert worden und auch optisch kaum erkennbar. Es liegen einige Kanten vor, die großteils parallel sind und eine plausible Orientierung aufweisen, aber es sind, wie sich bereits durch eine optische Analyse feststellen lässt, definitiv zu wenige, um die gesamte Szene rekonstruieren bzw. ein Höhenmodell flächendeckend verbessern zu können. Des Weiteren sind besonders im Kreuzriss auch einige Kanten zu erkennen, die offensichtlich falsch bestimmt wurden, da sie vertikal die gesamte Szene durchkreuzen.

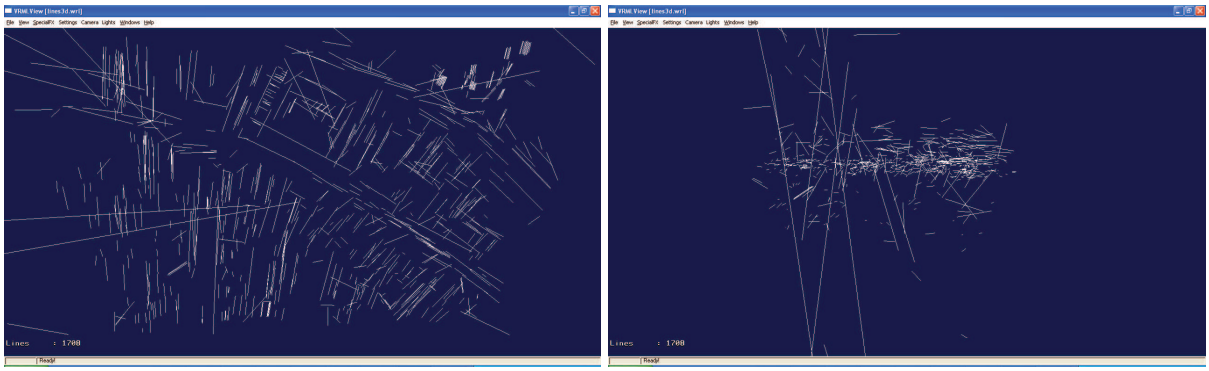

Abbildung 7.11: Grundriss (links) und Kreuzriss (rechts) der rekonstruierten 3D-Kanten. Als Input wurden die Luftbildkacheln des Flugstreifens C verwendet und die Standardparametereinstellungen gewählt.

Um solche Fehler auszuschließen wurde als Nächstes der Schwellwert für die Rückprojektionsbedingung vom Standardwert auf ein Fünftel reduziert (siehe Abbildung 7.12). Man erhält nun ein Ergebnis, bei dem diese Ausreißer fast vollständig eliminiert, allerdings auch viele korrekte Kanten ausgeschlossen werden. Es liegen nun insgesamt 628 rekonstruierte 3D-Kanten vor, im Vergleich zu 1708 im vorigen Beispiel. Daher kann man auch die Szene kaum noch erkennen, da sie nur sehr lückenhaft rekonstruiert wurde.

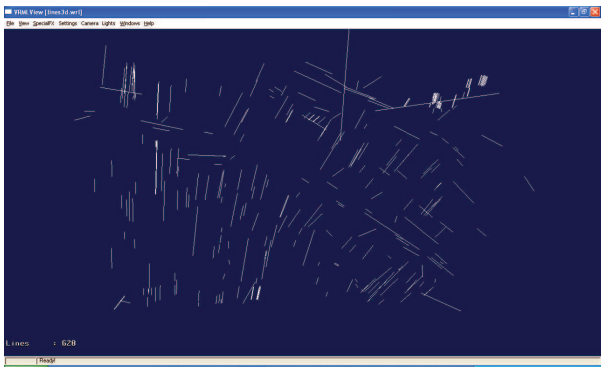

Abbildung 7.12: Grundriss der rekonstruierten 3D-Kanten aus den Luftbildkacheln des Flugstreifens C. Der Schwellwert für die Rückprojektionsbedingung wurde gegenüber dem Standardwert auf ein Fünftel reduziert.

Anschließend wurde ein Test mit dem sechsfachen Standardwert für den Schwellwert der Rückprojektionsbedingung durchgeführt (siehe Abbildung 7.13), um zu überprüfen, ob die Szene so vollständiger rekonstruiert werden kann. Das Resultat ähnelt großteils dem von Abbildung 7.11, nur dass geringfügig mehr 3D-Kanten (2186 im Vergleich zu 1708 vorhin) bestimmt wurden, von denen aber einige offensichtlich falsch sind. Die Szene kann damit auch nicht besser rekonstruiert werden.

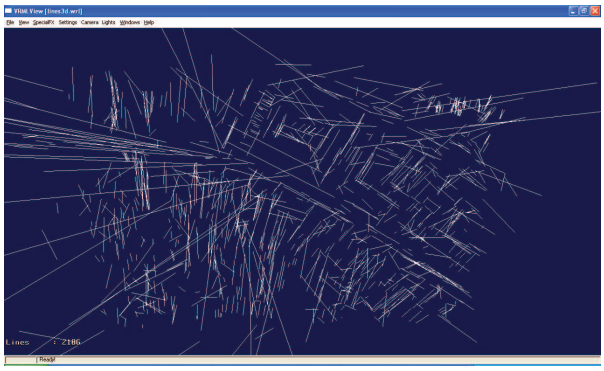

Abbildung 7.13: Grundriss der rekonstruierten 3D-Kanten aus den Luftbildkacheln des Flugstreifens C. Der Schwellwert für die Rückprojektionsbedingung wurde gegenüber dem Standardwert auf das Sechsfache erhöht.

Im Folgenden wurden die Luftbildkacheln von den gleich orientierten Flugstreifen A und C als Input verwendet (siehe Abbildung 7.14). Da die Standardparametereinstellungen bei den vorherigen Versuchen die besten Ergebnisse geliefert haben, werden diese wieder verwendet. Diesmal ist das Resultat allerdings nicht befriedigend, die räumliche Streuung der 3D-Kanten ist viel zu groß und entspricht keineswegs der Szene. Einzelne Linien sind von ihrer Orientierung und Position her plausibel, aber in Summe ist das Ergebnis unbrauchbar. Ein Grund für das Versagen des Algorithmus liegt wohl in der großen Anzahl von parallelen und ähnlichen Kanten. Diese wurden wahrscheinlich beim Matching falsch zugeordnet und führen zu falsch rekonstruierten 3D-Kanten, wie sie in diesem Beispiel häufig sichtbar sind. Auch andere Tests mit anderen Parametereinstellungen liefern keine besseren Ergebnisse.

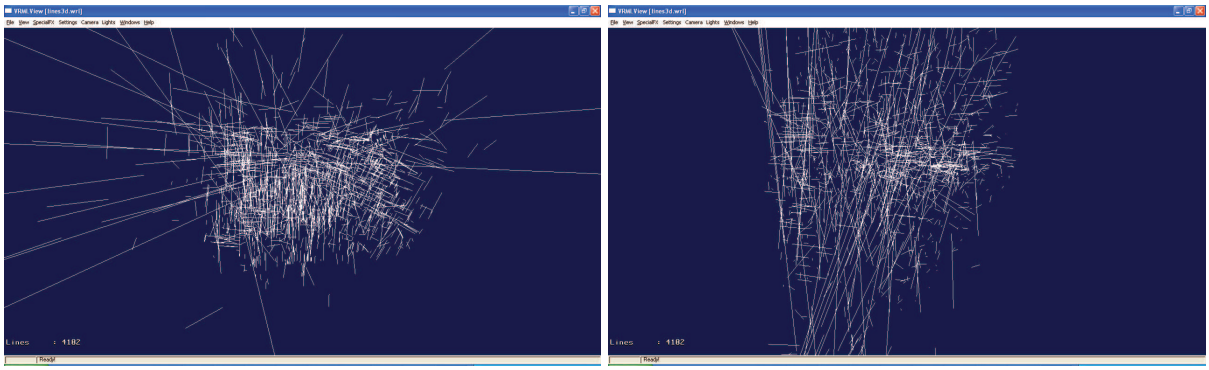

Abbildung 7.14: Grundriss (links) und Kreuzriss (rechts) der rekonstruierten 3D-Kanten. Als Input wurden die Luftbildkacheln der gleich orientierten Flugstreifen A und C verwendet und die Standardparametereinstellungen gewählt.

Die Implementierung wurde zuletzt auch mit allen erwähnten Luftbildkacheln der 3 Flugstreifen A, B und C getestet. Dabei konnte allerdings keine einzige 3D-Kante rekonstruiert werden. Somit liegt der Schluss nahe, dass der implementierte Algorithmus nicht sonderlich robust gegenüber Verdrehungen ist. Betrachtet man die der Implementierung beiliegenden Beispieldatensätze, welche auch Luftbilder beinhalten und allesamt gute Ergebnisse liefern, so fällt dabei auf, dass es sich immer um Bilder handelt, die kaum gegeneinander verdreht sind, zu einem sehr hohen Maß dieselbe Szene abbilden und diese auch noch eine deutlich leichter unterscheidbare Textur aufweist, als dies bei den verwendeten Datensätzen der Fall ist. Diese enthalten nämlich Szenen mit sich häufig wiederholenden Strukturen, wodurch viele räumlich naheliegende Liniensegmente detektiert werden, von denen, bedingt durch die koppelartige Bauweise im urbanen Raum, auch noch sehr viele parallel sind. Dies sind alles Herausforderungen, die ein Matching schwierig machen und den vorliegenden Algorithmus bei den verwendeten Luftbildern an seine Grenzen stoßen lässt.

8 Zusammenfassung und Ausblick

Wie gezeigt wurde, liefern die beiden getesteten Ansätze zum Matching von Liniensegmenten auch für stark überlappende (etwa 80% in und 60% quer zur Flugrichtung) Luftbilder keine ausreichend verlässlichen und genauen Ergebnisse, um damit flächendeckend die Traufenkanten für eine typische Szene im urbanen Raum zuverlässig und exakt rekonstruieren zu können. Ursachen für den mangelnden Erfolg der Methoden zum Matching von Liniensegmenten lassen sich in den Begebenheiten der Szene finden, konkret gemeint sind:

- die kaum unterscheidbare Textur
- und sich wiederholende Strukturen

Diese Faktoren machen ein erfolgreiches Matching unter Einbeziehung von ausschließlich spektraler Information de facto unmöglich. Daher wurde versucht, basierend auf der Verwendung der gegebenen Orientierung der Luftbilder, auch geometrische Information einzubeziehen. In der im Rahmen dieser Masterarbeit entwickelten Methode ist dies eine grob abgeschätzte Position der Liniensegmente, welche mithilfe eines groben, relativ ungenauen Digitalen Oberflächenmodells, bestimmt wurde. Darauf aufbauend wurde ein Suchfenster definiert, welches die Suche nach Korrespondenzen auf einige Matching-Kandidaten einschränkt. Dadurch wird die Wahrscheinlichkeit eines Fehlmatchings zwar deutlich reduziert, beim Matching dieser Kandidaten wird aber wieder nur spektrale Information herangezogen, wodurch die zuvor genannten Ursachen, trotz des Suchfensters, zu relativ vielen Missmatches und damit zu einem unverlässlichen Ergebnis führen.

Auch durch die Einbeziehung weiterer geometrischer Information, wie es bei der Methode von Schmid und Zisserman durch die Einführung einer Epipolar- und Rückprojektionsbedingung geschieht, wurden keine besseren Resultate erzielt. Hinzu kommt, dass dieser Ansatz offensichtlich nicht robust gegenüber größeren Verdrehungen ist, was die Anwendungen auf Luftbilder entgegengesetzter Flugstreifen unmöglich macht.

Natürlich gibt es etliche Möglichkeiten, die bestehenden Methoden zu optimieren, sei es durch eine an die Szene angepasste und feinere Unterteilung des HSV-Farbraums im Zuge der Berechnung der Farbprofilhistogramme, welche bei der entwickelten Methode in Anlehnung an den Ansatz von Bay et al. als Deskriptor fungieren, oder sei es, indem nur die zentralen Bereiche der Luftbilder verwendet werden, wodurch Probleme mit auftretenden Relief-verschiebungen reduziert werden. Es ist allerdings fraglich, ob derartige Optimierungen ausreichend wärem, um zu dem gewünschten Ergebnis, nämlich der exakten Rekonstruktion von Szenen vor allem im urbanen Raum und mit dem Hauptaugenmerk auf Traufen, zu kommen, welches wiederum die Verbesserung von Höhenmodellen und damit die Generierung genauerer Orthophotos ermöglichen würde, oder ob es hierzu nicht anderer Methoden bedarf. Untersuchungen in diese Richtung würden den Rahmen dieser Arbeit sprengen und wurden deshalb nicht in Angriff genommen, sie könnten aber Gegenstand nachfolgender Arbeiten sein.

9 Literaturverzeichnis

H. Bay, V. Ferrari, L. V. Gool. Wide-baseline stereo matching with line segments. International Conference on Computer Vision & Pattern Recognition, 329–336, 2005.

J. Canny. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 679–698, 1986.

N. Chehata, M. P. Desseilligny, F. Jung, G. Stamon. Extraction of 3D primitives from stereopairs of satellite images for automatic reconstruction of buildings. IAPR Workshop on Machine Vision Applications, 636–639, 2002.

B. Fan, F. Wu, Z. Hu. Line matching leveraged by point correspondences. International Conference on Computer Vision & Pattern Recognition, 390–397, 2010.

A. F. Habib, R. Zhai, C. Kim. Generation of complex polyhedral building models by integrating stereo aerial imagery and LIDAR data. Photogrammetric Engineering & Remote Sensing 76 (5), 609–623, 2010.

G. Heidelmeyer. Methoden zur Analyse, Bewertung und Modifikation digitaler Höheninformationen für Luftfahrtanwendungen. Diss., TU Darmstadt, 2008.

F. Jung, N. Paparoditis. Extracting 3D free-form surface boundaries of manmade objects from multiple calibrated images: a robust, accurate and high resolving power Edgel matching and chaining approach. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 34, 39–44, 2003.

T. Kersten. Das Projekt swissphoto - Flächendeckende und aktuelle Geodaten als Grundlage für viele GIS-Anwendungen in der Schweiz. GIS Geo-Informations-Systeme 6/96, 12–19, 1996.

K. Kraus. Photogrammetrie Bd. 3, Dümmler, Köln, 2000.

K. Kraus. Photogrammetrie Bd. 1, 7. Aufl., Walter de Gruyter, Berlin, 2004.

D. Kulbach. Systematischer Vergleich von Verfahren zur überwachten Klassifikation von multispektralen Luftbildern mit Bodenauflösungen im Meterbereich. Dipl.-Arb., Univ. Hamburg, 1997.

G. Navratil. Ausgleichungsrechnung II (Skriptum). TU Wien, Department für Geodäsie und Geoinformation, 2012.

M. Ø. Nielsen. True orthophoto generation. Diss., TU of Denmark, Lyngby, 2004.

N. Pfeifer. Oberflächenmodelle aus Laserdaten. VGI Österreichische Zeitschrift für Vermessung & Geoinformation 91/4, 243−253, 2003.

C. Schmid, A. Zisserman. Automatic line matching across views. International Conference on Computer Vision & Pattern Recognition, 666–671, 1997.

F. Taillandier, R. Deriche. Reconstruction of 3d linear primitives from multiple views for urban areas modelisation. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 34, 267–272, 2002.

R. G. Von Gioi, J. Jakubowicz, J. M. Morel, G. Randall. LSD: a line segment detector. Image Processing On Line 2, 33−55, 2012.

L. Wang, U. Neumann, S. You. Wide-baseline image matching using line signatures. International Conference on Computer Vision, 2009.

Z. Wang, F. Wu, Z. Hu. MSLD: A robust descriptor for line matching. Pattern Recognition 42, 889–896, 2009.

T. Werner, A. Zisserman. New techniques for automated architectural reconstruction from photographs. Computer Vision–ECCV. Springer, Berlin-Heidelberg, 541–555, 2002.

L. Zhang, R. Koch. Line matching using appearance similarities and geometric constraints. Lecture Notes in Computer Science 7476, 236–245, 2012.

Internet-Literaturangaben:

HSV-Farbraum (13.01.2015):
http://upload.wikimedia.org/wikipedia/commons/0/0d/HSV_color_solid_cylinder_alpha_lowgamma.png

Epipolargeometrie (15.01.2015):
https://upload.wikimedia.org/wikipedia/commons/3/33/Epipolargeometrie3.svg

Trifokalgeometrie (15.01.2015):
https://upload.wikimedia.org/wikipedia/commons/f/fb/Trifokalgeometrie.svg

Ultracam (31.01.2015):
https://download.microsoft.com/download/7/4/3/743EFD09-258B-4BFA-8D56-3148C60DD137/UCAMTechnicalDocuments/UltraCamX-Specs.pdf