

Punktkorrespondenzen in Bildpaaren aus projektiven und radiometrischen Invarianzen

Zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften
(Dr.-Ing.)

bei der Fakultät für
Bauingenieur-, Geo- und Umweltwissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte
DISSERTATION

von

Dipl.-Ing. Simon Schuffert
aus Hannover

Datum der mündlichen Prüfung: 13. Juli 2015
Referent: Prof. Dr. Stefan Hinz
Korreferent: apl. Prof. Dr. Franz Rottensteiner

Kurzfassung

Eine fundamentale Voraussetzung für sehr viele Anwendungen in der Photogrammetrie und in der Computer Vision ist es, identische Punkte eines abgebildeten Objektes in zwei sich überlappenden Bildern zu finden.

Diese Fähigkeit der automatischen Zuordnung gleicher Objekte, welche unter stark unterschiedlichen Blickwinkeln mit unterschiedlichen Kameras und zu unterschiedlichen Zeitpunkten aufgenommen wurden, ist bisher aus verschiedenen Gründen nur sehr eingeschränkt gelöst. Die vorliegende Arbeit beschreibt mehrere Lösungswege für Punktzuordnungen extremer Blickwinkelveränderungen, die zudem unabhängig von Beleuchtungsveränderungen sind.

Zum Auffinden von Punktkorrespondenzen ist das Identifizieren homologer Objektpunkte in unterschiedlichen Bildern notwendig. Die Zuordnungen von Pixeln im ersten Bild zu ihren Entsprechungen im zweiten Bild, geschehen über gleiche Merkmale. Dafür werden bisher meistens lokale Merkmale in beiden Bildern beschrieben, die dann miteinander verglichen werden. Aktuelle Featuredeskriptoren wie Harris [Harris & Stephens, 1988], Harris-Affine [Mikolajczyk & Schmid, 2002], SIFT [Lowe, 2004], SURF [Bay *et al.*, 2006], FAST [Rosten & Drummond, 2006], MSER [Matas *et al.*, 2004], ORB [Rublee *et al.*, 2011], BRISK [Leutenegger *et al.*, 2011] oder ASIFT [Yu & Morel, 2009] beschreiben punktbezogene oder blobbezogene Merkmale, die für einen automatischen Algorithmus gut zuzuordnen sind. Diese Merkmalsdeskriptoren beschreiben die kleinräumige Textur um ein Pixel herum und funktionieren hervorragend bei geringen Änderungen des Blickwinkels gegenüber dem Objekt, wenn die Aufnahmezeitpunkte sehr nah beieinander liegen und auch wenn für beide Bilder der gleiche Sensortyp benutzt wurde. Bei stark unterschiedlichen Blickpunkten auf ein dreidimensionales Objekt können diese Featuredeskriptoren hingegen keine korrekten Ergebnisse liefern. Deutliche Änderungen in der Radiometrie des Bildes aufgrund unterschiedlicher Beleuchtungen oder auch veränderte Schattenverläufe aufgrund unterschiedlicher Tageszeitpunkte verhindern erfolgreiche Punktzuordnungen. Auch der Ansatz des *Bag of visual words*, der eine allgemeinere Bildbeschreibung liefert, ist aufgrund der verwendeten Algorithmen lediglich invariant gegenüber affinen Transformationen [Bronstein *et al.*, 2011; Ren *et al.*, 2014; Sivic & Zisserman, 2003].

Der zentrale Ansatz und die Idee der vorliegenden Arbeit zur Lösung dieser Probleme ist folgender: Beide Bilder enthalten zweidimensionale Abbildungen von identischen dreidimensionalen Objekten, die mit einer Zentralprojektion auf die Bildebene

abgebildet wurden. Beiden Bildern gemeinsam ist somit, dass Ebenen eines Objektes mit einer projektiven Transformation abgebildet wurden. Zum Auffinden von identischen Punkten werden gleiche Eigenschaften in beiden Bildern gesucht, die sich durch eine projektive Transformation nicht verändern und somit projektiv invariant sind. Beide Bilder werden daher nur mit Hilfe von Eigenschaften beschrieben, die projektiv invariant sind. Anschließend werden diese Bildbeschreibungen miteinander verglichen und die damit verbundenen identischen Punkte einander zugeordnet. Diese projektiv invarianten Eigenschaften sind im Einzelnen: Das Doppelverhältnis von vier Punkten auf einer projektiven Geraden, die zwei Invariante I_1 und I_2 von fünf Punkten auf einer projektiven Ebene, die Konnektivität von Punkten, die Inzidenzen von Punkten sowie die Beibehaltung gerader Linien. Um unabhängig von der Variabilität der Radiometrie zu sein, werden die Punktzuordnungen ausschließlich anhand von Gradientenamplitudensprüngen beschrieben.

Die Verwendung einzelner projektiv invarianter Eigenschaften zum Bestätigen identischer Bildpunkte ist grundsätzlich nicht neu. Die Neuheit und das wesentliche Alleinstellungsmerkmal der Arbeit bestehen darin, dass die Beschreibung eines Bildes mittels mehrerer kombinierter projektiv invarianter Eigenschaften abgegeben wird. Diese Kombination führt zu völlig neuen Möglichkeiten der Bildkorrespondenzsuche. Ein weiterer Vorteil dieses Lösungsansatzes ist, dass dieser auch auf Bilder anwendbar ist, die mit einfacheren Transformationen abgebildet wurden, da die projektiv invarianten Eigenschaften auch für diese Transformationen gelten wie beispielsweise eine Translation, eine Rotation, eine Skalierung oder eine affine Abbildung. Eine Lösung des Zuordnungsproblems für projektive Abbildungen löst daher auch das Zuordnungsproblem bei untergeordneten Transformationen.

Die Ergebnisse der Arbeit und die durchgeführten Experimente zeigen, dass mittels kombinierter projektiv invarianter Merkmale Punktzuordnungen gefunden werden können, welche mit bisherigen Verfahren nicht möglich waren.

Abstract

A fundamental requirement for many applications in photogrammetry and computer vision is to find identical points of an object shown in two overlapping images. This is an ability that is relatively easy for humans. But in automatic image analysis this is only solved under limited conditions.

The ability of automatic assignment of the same objects, which were taken from very different angles with different cameras and at different times, is a challenging task for various reasons. This thesis describes a number of solutions for point assignments under extreme viewing angle changes, which are also independent of illumination changes. For finding point correspondences it is necessary to find the same object points in different images. The assignments of pixels in the first image to their counterparts in the second image, are based on the same features. Very often local characteristics are described in the two images, which are then compared with each other. Current feature descriptors as Harris [Harris & Stephens, 1988], Harris-Affine [Mikolajczyk & Schmid, 2002], SIFT [Lowe, 2004], SURF [Bay *et al.*, 2006], FAST [Rosten & Drummond, 2006], MSER [Matas *et al.*, 2004], ORB [Rublee *et al.*, 2011], BRISK [Leutenegger *et al.*, 2011] or ASIFT [Yu & Morel, 2009] describe characteristics of points or blobs that are easy to find and to assign for an algorithm.

These feature descriptors describe a small area around the pixel and they work excellent if the change of the viewing angle is small and if the illumination does not change very much and the used sensor is the same. However when taking very different viewpoints on an object, these feature descriptors deliver erroneous results. Also significant changes in the radiometry of the image due to different illumination conditions or changing shadow casts at different times prevent successful point assignments.

Also the approach of the *Bag of Visual Words*, which provides a general image description is, due to the algorithms used, only invariant to affine transformations [Bronstein *et al.*, 2011; Ren *et al.*, 2014; Sivic & Zisserman, 2003].

The central approach and the idea behind this work is to solve these problems as follows: Both images contain two-dimensional mappings of the same three-dimensional objects that have been mapped with a central projection onto the image plane. Therefore both mappings inhere the same characteristics. To locate identical points, same properties are searched for in both images, which do not change under a projective transformation and thus are projective invariant. Both images will therefore be

described exclusively with the help of properties that are projective invariant. Subsequently, these image descriptions are compared and assigned to the associated identical points. These projective invariant properties are as follows: The cross ratio of four points on a line, the two invariants I_1 and I_2 of five points on a projective plane, the connectivity of points, the incidences of points and the preservation of straight lines. To be independent of the variability of radiometry, the point assignments are exclusively described based on gradient amplitude changes.

The use of individual projective invariant properties to find identical pixels is not new. The novelty and the main unique contribution of this work is that the description of an image is based on a variety of combined projective invariant properties. This combination leads to completely new ways of image correspondence search. Another advantage of this approach is that this is also applicable to images that have been mapped with simpler transformations, since the projective invariant properties also apply to these transformations such as translation, rotation, scaling or an affine transformation. Therefore, a solution of the assignment problem for projective transformations solves the assignment problem with subordinate transformations as well.

The results of the work and the experiments carried out show that by means of combined projective invariant features point assignments can be found, which could not be performed by previous methods.

Ich, Simon Schuffert, erkläre, dass ich die vorliegende Arbeit, abgesehen von der Benutzung der von mir vollständig und genau bezeichneten Hilfsmittel, selbstständig verfasst habe und die Grundsätze des Karlsruher Instituts für Technologie (KIT) zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet habe.

Karlsruhe, den 10.06.2015

Simon Schuffert

Danksagung

Schlage die Trommel und fürchte dich nicht heißt es bei Heinrich Heine [Heine, 1844]. Viele haben mir beigebracht zu trommeln, indem sie mir durch fachliche Diskussionen, durch ihre Unterstützung, durch ihre Hilfsbereitschaft und auch menschlich geholfen oder als Vorbilder gedient haben.

Ich danke Prof. Dr. Stefan Hinz für die ausgezeichnete Betreuung dieser Arbeit, für die fachlichen Diskussionen, die wir geführt haben, für seine vorbildliche Besonnenheit und für die optimalen Bedingungen, die ich in meiner Zeit am Institut für Photogrammetrie und Fernerkundung (IPF) hatte. Ich danke Prof. Dr. Franz Rottensteiner für die erstklassige Betreuung als Korreferent und für die hilfreichen Diskussionen und Anregungen, die in diese Arbeit eingeflossen sind. Ich danke Prof. Dr. Hans-Peter Bähr für die vielen auf langjährigen Erfahrungen beruhenden praxisbezogenen Anregungen und für die hilfreichen Diskussionen. Ich danke Dr. Sven Wursthorn für jahrelange fachliche Diskussionen und technische Unterstützung. Ich danke Dr. Thomas Vögtle für die vielen fachlichen Diskussionen, für die tolle Teamarbeit und seine vorbildliche sachliche Herangehensweise an große Herausforderungen. Ich danke Dr. Fadwa Alshawaf, Dr. Martin Weinmann, Steffen Urban, Konrad Ringle, Anna Giovanna und Rosmarie Blomley für die vielen fachlichen Diskussionen und für deren Hilfsbereitschaft. Ich danke Ilse Engelmann und Heike Birkel für die freundliche Unterstützung und Zusammenarbeit. Dem gesamten Team des IPF danke ich für die vielen Ratschläge und Unterstützungen, die ich erhalten habe.

Dr. Christian Lucas und Dr. Andreas Schenk, den Koordinatoren der Graduiertenschule GRACE, danke ich für ihr Engagement, durch das mir wichtige Auslandsaufenthalte und Fortbildungen ermöglicht wurden. Ich danke James Eddy von Bluesky International Limited für die freundliche Unterstützung und die Bereitstellung der Luftbilder, welche bei der Lösungsfindung ausschlaggebend waren.

Ohne Fleiß kein Preis! Daran haben mich meine Eltern Eberhard und Esther Schuffert früher oft erinnert. Ich danke ihnen ganz besonders dafür, dass sie mich immer voll unterstützt haben. Ich danke Dr. Joachim Wiesel für sein Vertrauen in mich und für die vorbildliche zielführende Teamarbeit bei der Durchführung umfangreicher Projekte. Ich danke Prof. Dr. Arivaldo Leão de Amorim für die überwältigende Gastfreundschaft in Salvador da Bahia und für die gemeinsame Projektarbeit an mehreren Orten im Bundesstaat Bahia in Brasilien. Ich danke Torsten Lauer für die vielen Tipps bei handwerklichen Projekten und für seine Hilfsbereitschaft. Ich danke Martin Rittershofer dafür, dass er es mir sowohl auf unseren Klettersteig-Touren in den Alpen

als auch bei beruflichen Zielsetzungen jedesmal schmackhaft macht, höhere Gipfel als die bisherigen zu besteigen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Zielsetzung und wissenschaftliche Beiträge dieser Arbeit	3
1.3	Gliederung der Arbeit	4
2	Stand der Forschung und Einbettung in den wissenschaftlichen Kontext	5
2.1	Vergleiche und Literaturrecherche	5
2.1.1	Zentralprojektion und Epipolargeometrie	5
2.1.2	Aktuelle Konzepte und Algorithmen für Punktzuordnungen	7
2.1.3	Einschränkungen aktueller Konzepte und deren Folgen	8
2.2	Projektive Invarianzen	12
2.2.1	Invariante Kenngrößen bei unterschiedlichen Transformationen	12
2.2.2	Aktuelle Methoden basierend auf projektiven Invarianzen	16
2.3	Radiometrische Invarianzen	20
2.3.1	Herausforderungen radiometrisch veränderter Bildinformation	20
2.3.2	Aktuelle Methoden basierend auf radiometrischen Invarianzen	21
2.4	Abgeleitete Schlüsse für den neuen Bildzuordnungs-Ansatz	22
2.4.1	Neue Beiträge dieser Arbeit	22
2.4.2	Unterschiede zu bestehenden Arbeiten	23
3	Grundlagen und Konzeption	24
3.1	Konzeption projektiver Invarianzen	24
3.1.1	Gerade Linien	24
3.1.2	Inzidenzen	25
3.1.3	Projektiv planar invariante Merkmale	27
3.1.4	Konnektivität	28
3.2	Konzeption radiometrischer Invarianzen	30
3.2.1	Ansprüche an die Merkmale	30
3.2.2	Kanten als Inzidenzen	32
3.2.3	Eckpunkte als Inzidenzen	32
3.3	Anwendung bekannter projektiv invarianter Flächenmerkmale	34
3.3.1	Flächenmerkmale mit Hilfe von P^2 -Invarianten	34
3.3.2	Spezifische und eingeschränkte Verbesserungsansätze	36
4	Neues Konzept der Zuordnung virtueller Linienmerkmale	38
4.1	Grundidee	38

Inhaltsverzeichnis

4.2	Merkmalsbeschreibung bei Linienmerkmalen	40
4.2.1	Eckpunkte und Kanten als Eingangsdaten	40
4.2.2	Erstellung der Linienmerkmale	40
4.3	Zuordnung von Linienmerkmalen	42
4.3.1	Vorgehensweise	42
4.3.2	Selektion beteiligter Linienfeatures	42
4.3.3	Vergleich von Linienfeatures anhand von Doppelverhältnissen	44
4.3.4	Nachbarschaftsreihenfolge testen	44
4.3.5	Ausgabe möglicher Punktkorrespondenzen	45
5	Neues Konzept der kombinierten Linien- und Flächenmerkmale	46
5.1	Virtuelle Linienmerkmale für die initialen Hypothesen	46
5.2	Kombinierte Linien- und Flächenmerkmale	47
5.2.1	Test aller Kandidatenpaare	47
5.2.2	Plausibilitätstests	47
5.2.3	Logische Reduzierung der Kandidatenliste	48
5.2.4	Abschließende Überprüfung durch virtuelle Linienfeatures . . .	49
6	Evaluation	52
6.1	Synthetische Testdaten	52
6.2	Sensitivitätsanalysen	54
6.2.1	Eckpunkt-Features und Blickwinkelveränderungen	54
6.2.2	P^2 -Invariante I_1 und I_2 und Blickwinkelveränderungen	55
6.2.3	Unterschiedliche perspektivische Deformationen	57
6.2.4	Exemplarischer Test von unterschiedlichen Radiometrien	60
6.2.5	Punktkorrespondenzen bei Untermengen	62
6.2.6	Fehlerhafte Eckpunkte	63
6.2.7	Doppelverhältnisse in den Linienfeatures	64
6.2.8	Zusammenspiel von Nachbarschaftsgröße und P^2 -Invariante . .	65
6.3	Übertragbarkeit der Verfahren und Realbilder	74
6.4	Mehrere Ebenen	74
6.5	Ergebnisvergleich mit anderen Verfahren	78
6.5.1	Ergebnisvergleich: Blobbezogene Verfahren	78
6.5.2	Ergebnisvergleich: Projektiv invariante Verfahren	79
6.5.3	Konzeptionell bedingte Unterschiede	80
7	Zusammenfassung	82
8	Ausblick und Vorschläge für zukünftige Arbeiten	84
Anhang		86

Max Planck:

Eine neue wissenschaftliche Wahrheit pfl egt sich nicht in der Weise durchzusetzen, da ß ihre Gegner überzeugt werden und sich als belehrt erklären, sondern vielmehr dadurch, da ß die Gegner allmählich aussterben und da ß die heranwachsende Generation von vornherein mit der Wahrheit vertraut gemacht ist.

Wissenschaftliche Selbstbiographie
Johann Ambrosius Barth Verlag, Leipzig, S.22, 1948

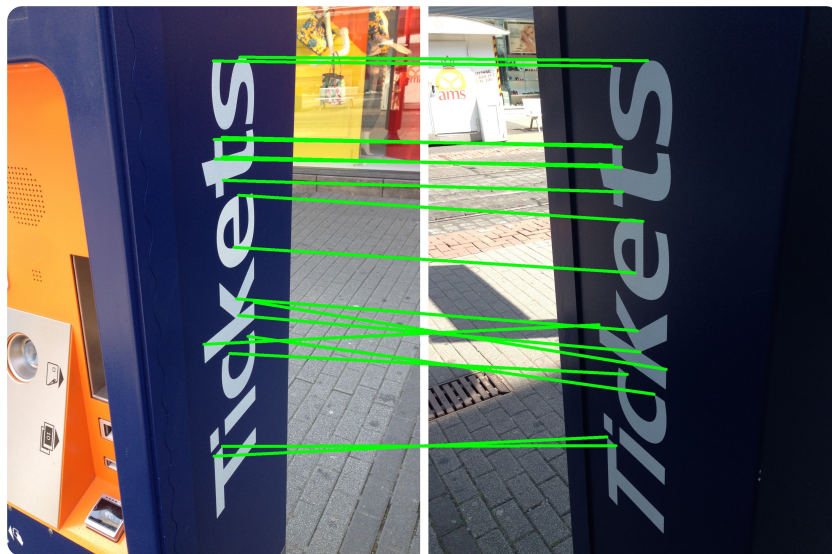


Abbildung 1: Erfolgreich detektierte Punktkorrespondenzen aus projektiven und radiometrischen Invarianzen in einem Stereobildpaar mit extrem weiter Basis.

1 Einleitung

1.1 Motivation

Ein Mensch kann relativ einfach identische Objekte in zwei Bildern wiederfinden. Wenn er sich beispielsweise bei einem Rundflug in einem Flugzeug über einer Stadt bei sonnigem Wetter mit seiner Kamera ein Bild von einem Marktplatz aus der Luft aufnimmt und denselben Marktplatz an einem anderen Tag bei Regen vom Boden aus fotografiert, so kann er eindeutig dieselben Häuserfassaden, dieselben Bäume und dieselben Springbrunnen in beiden Bildern einander zuordnen. Das trifft sogar dann zu, wenn beide Bilder nicht bloß zu unterschiedlichen Zeitpunkten sondern auch mit unterschiedlichen Kameras und eventuell sogar mit unterschiedlichen Sensoren aufgenommen wurden. Auch Schattenwürfe sowie durch verschiedene Lichtquellen veränderte Farben oder sich im Laufe der Zeit verändernde Oberflächen halten den Menschen nicht davon ab, schnell und intuitiv korrespondierende Strukturen in beiden Bildern zu finden. Mit Hilfe solcher identischen Objektstrukturen können beispielsweise dreidimensionale Rekonstruktionen der Szene vorgenommen werden, oder einfach nur die relative Lage der beiden Bilder zueinander bestimmt werden, was Grundvoraussetzung für viele weitere Anwendungen ist.

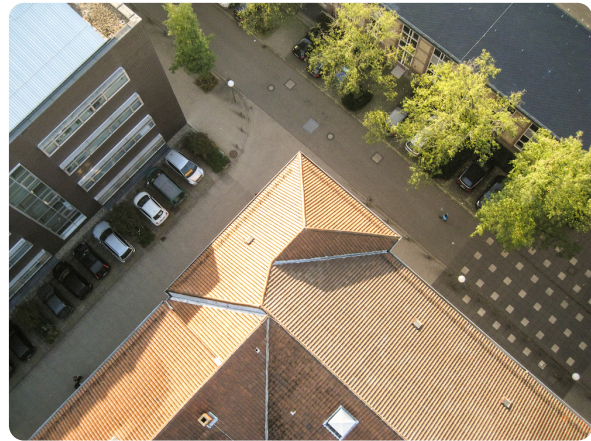
Im Bereich der automatischen Verarbeitung von Bildern in der Photogrammetrie und in der Computer Vision stellen genau diese genannten Bedingungen ein hochkomplexes Problem dar. Das Auffinden von Punktkorrespondenzen in zwei sich überlappenden Bildern wurde bisher vor allem auf das Problem zurückgeführt, identische lokale Bereiche in beiden Bildern zu finden. Lokale Merkmale haben sich bisher als besonders geeignet erwiesen, unter anderem weil sie einfach und mit wenig Rechenaufwand zu erstellen sind und bisherige Anforderungen hervorragend lösen konnten. Hingegen Merkmale, die größere Bereiche eines Bildes beschreiben, wie Linien- oder Flächenmerkmale, sind schwieriger zu berechnen oder zuzuordnen. Aktuell benutzte lokale Merkmalsdetektoren sind nach [Mikolajczyk *et al.*, 2005] und [Weinmann, 2013] beispielsweise: Harris [Harris & Stephens, 1988], Harris-Affine [Mikolajczyk & Schmid, 2002], SIFT [Lowe, 2004], SURF [Bay *et al.*, 2006], FAST [Rosten & Drummond, 2006], MSER [Matas *et al.*, 2004], ORB [Rublee *et al.*, 2011], BRISK [Leutenegger *et al.*, 2011] oder ASIFT [Yu & Morel, 2009]. Diese basieren auf lokalen Merkmalen wie z.B. Ecken oder Blobs und sind stark von der Textur und den Intensitätswerten der einzelnen Kanäle eines Bildes am Punkt und seiner direkten Umgebung abhängig. Auch wenn alle Verfahren in ihren Anwendungsgebieten jeweils

1 Einleitung

hervorragende Ergebnisse liefern, wird bei allen Techniken ein wesentlicher Aspekt der beschriebenen Bilder außer Acht gelassen, nämlich dass bei der Erstellung eines Bildes jeweils eine zentralperspektivische Abbildung der dreidimensionalen Objekte in eine zweidimensionale Bildebene vorgenommen wurde [Dürer, 1525]. Versucht ein Algorithmus ein Merkmal aus einem Bild in einem anderen Bild wiederzufinden, so muss dieses Merkmal auch invariant gegenüber der jeweiligen Projektion sein, die notwendig ist, den Teilbereich des einen Bildes in den des anderen zu transformieren, siehe Beispiel in Abbildung 1.1. Um Merkmale zuverlässig wiederzufinden bedeutet das, dass sie projektiv invariant sein müssen, wenn sie zum Vergleich in beiden Bildern herangezogen werden. Aufgrund ihrer Konzeption sind die bisherigen Algorithmen jedoch maximal gegenüber affinen Transformationen invariant [Mikolajczyk *et al.*, 2005]. Auch neuere Ansätze im Bereich von Linienmerkmalen wie *Descriptor-Nets* [von Hundelshausen & Sukthankar, 2012] sind aufgrund der benutzten Funktionen ebenfalls nur invariant gegenüber affinen Transformationen (siehe Kapitel 4).



(a) Gebäude vom Boden aus aufgenommen.



(b) Dasselbe Gebäude wie in Bild a zu einem früheren Zeitpunkt und aus der Vogelperspektive von einer fliegenden Kleinplattform aus aufgenommen (links oben im Bild).

Abbildung 1.1: Ein Beispiel für einen stark veränderten Blickwinkel. Die starke perspektivische Transformation zwischen Teilbereichen der Bilder erschwert eine automatische Suche nach identischen Punkten in beiden Bildern.

Eine weitere wesentliche Rolle bei der Merkmalsbeschreibung spielt auch die Radiometrie. Da sich der Lichteinfall auf das gleiche Objekt zwischen beiden Aufnahmezeitpunkten stark verändern kann, ergeben sich oftmals stark unterschiedliche radiometrische Beschreibungen der ursprünglich gleichen Oberfläche, siehe Abbildung 1.2. Ein Verfahren, das allgemein anwendbar sein soll, muss daher auch invariant gegenüber radiometrischen Veränderungen sein. Derzeit stellen die oben beschriebenen Problemstellungen große Herausforderungen dar.



(a) Luftbildaufnahme am Tag.



(b) Luftbildaufnahme bei Nacht.

Abbildung 1.2: Ein Beispiel für veränderte Radiometrie: Die zwei Luftbildaufnahmen bilden das gleiche Gebiet in der Stadt Leicester in England ab. Durch unterschiedliche Lichtquellen und Schattenwürfe ist die Textur in beiden Bildern stark unterschiedlich, was eine automatische Suche nach Punktkorrespondenzen erschwert. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)

1.2 Zielsetzung und wissenschaftliche Beiträge dieser Arbeit

Im Bereich der automatischen Koregistrierung von beliebigen Bildern, welche dieselben Objekte darstellen, spielen projektiv invariante Merkmale bisher nur eine untergeordnete Rolle. Ihre Erstellung und der Vergleich der Merkmale verursacht nach aktuellem Stand der Technik einen extrem hohen Rechenaufwand (siehe Kapitel 3.3) im Gegensatz zu der Erstellung vergleichsweise einfacher lokaler Merkmale.

Die Neuheit dieser Arbeit besteht darin, dass mehrere neue Lösungswege konzipiert werden, mit denen sowohl projektiv als auch radiometrisch invariante Merkmale beschrieben werden. Diese Merkmale werden dann anschließend miteinander verglichen, um identische Punkte in beiden Bildern zu finden. Der neue Lösungsweg dieser Arbeit ist es, den Inhalt eines Bildes nur mit Hilfe von kombinierten projektiv invarianten Merkmalen zu beschreiben. Anschließend wird diese Bildbeschreibung mit der Beschreibung anderer Bilder verglichen, die auf dieselbe Art charakterisiert wurden.

Zusammengefasst sind die Anforderungen an den in dieser Arbeit neu zu entwickelnden Ansatz zum Auffinden von Punktkorrespondenzen in Stereobildern folgen-

1 Einleitung

de:

1. Der neue Ansatz soll projektiv invariante Bildzuordnungen ermöglichen, was extrem starke Blickwinkelunterschiede zwischen den beiden Bildern beinhaltet.
2. Der neue Ansatz soll es ermöglichen, Bilder mit deutlich unterschiedlichen Kamerakonstanten, was bei Blickwinkelveränderungen in unterschiedlichen optischen Verkürzungen resultiert, zuzuordnen.
3. Der neue Ansatz soll radiometrisch invariantes Matching ermöglichen, wodurch Beleuchtungsunterschiede oder unterschiedliche Sensortypen in beiden Bildern handhabbar werden.
4. Das neue Konzept soll übertragbar und allgemein anwendbar sein für unterschiedliche Anwendungsbereiche.
5. Der neue Ansatz soll auf Teilausschnitte von Bildern anwendbar sein, was auch Verdeckungen mit einschließt.
6. Gegenüber bestehenden Lösungsansätzen, welche projektiv invariante Einzelpunktzuordnungen unter einem enormem Rechenaufwand anbieten, soll eine signifikante Reduzierung von Rechenschritten erreicht werden.

1.3 Gliederung der Arbeit

Die folgenden sieben Kapitel gliedern sich folgendermaßen. In Kapitel 2 wird der Stand der Forschung auf dem Thema der Suche nach Punktkorrespondenzen erörtert, und basierend darauf werden die Schlüsse für den neuen Lösungsansatz dieser Arbeit abgeleitet. Die Grundlagen, die für diesen neuen Lösungsansatz vorausgesetzt werden, werden im darauffolgenden Kapitel 3 erklärt. Danach folgt der Teil der Arbeit, in dem die neu entwickelten Konzepte vorgestellt werden. In Kapitel 4 wird das neue Konzept der Zuordnung von virtuellen Linienmerkmalen vorgestellt. Im Anschluss daran werden in Kapitel 5 die ebenfalls neue Kombinationen aus virtuellen Linien- und Flächenmerkmalen vorgestellt. Danach werden die vorgestellten neuen Algorithmen evaluiert (Kapitel 6). Zum Abschluss wird in Kapitel 7 der wissenschaftliche Erkenntnisgewinn diskutiert und in Kapitel 8 werden Ideen für neue zukünftige Forschungsarbeiten, die aus den Ergebnissen dieser Arbeit entstanden sind, vorgestellt.

2 Stand der Forschung und Einbettung in den wissenschaftlichen Kontext

2.1 Vergleiche und Literaturrecherche

2.1.1 Zentralprojektion und Epipolargeometrie

Das meistbenutzte Abbildungsmodell der *Photogrammetrie* und der *Computer Vision* ist die zentralprojektive Abbildung. Mathematisch gesehen definiert diese, wie der projektive Raum mit Hilfe der linearen Algebra in homogenen Koordinaten abgebildet wird. Im Bereich der Photogrammetrie beschreibt die Zentralprojektion, wie mit Hilfe einer Kamera die Objekte der dreidimensionalen Welt in die Bildebene abgebildet werden. Diese Abbildung folgt den Regeln der projektiven Geometrie. Im einfachsten Fall wird dem Kameramodell eine Lochkamera zugrunde gelegt. Die Abbildungsvorschriften zwischen den Objektpunkten im dreidimensionalen Raum und den homogenen Koordinaten des Kamerakoordinatensystems sind durch die Kollinearitätsgleichungen [Albertz & Wiggenhagen, 2009; Luhmann, 2010] gegeben, siehe Gleichungen 2.1 und 2.2.

$$x'_P - x'_0 = -c_k \frac{r_{11}(X_P - X_0) + r_{21}(Y_P - Y_0) + r_{31}(Z_P - Z_0)}{r_{13}(X_P - X_0) + r_{23}(Y_P - Y_0) + r_{33}(Z_P - Z_0)} \quad (2.1)$$

$$y'_P - y'_0 = -c_k \frac{r_{11}(X_P - X_0) + r_{21}(Y_P - Y_0) + r_{31}(Z_P - Z_0)}{r_{13}(X_P - X_0) + r_{23}(Y_P - Y_0) + r_{33}(Z_P - Z_0)} \quad (2.2)$$

x'_P, y'_P :	Bildkoordinaten des Objektpunktes
x'_0, y'_0 :	Bildhauptpunkt
c_k :	Kamerakonstante
r_{ij} :	Parameter der Drehmatrix
X_P, Y_P, Z_P :	Koordinaten des Objektpunktes
X_0, Y_0, Z_0 :	Koordinaten des Projektionszentrums O

Die Suche nach identischen Punkten in mehreren Bildern ist auch deshalb wichtig, weil daraus die innere und die äußere Orientierung der Kamera errechnet werden können. Die Zusammenhänge werden über die Epipolargeometrie ausgedrückt und

stellen sich wie folgt dar (nach [Albertz & Wiggenhagen, 2009] S.214-215, verändert):

Die Epipolargeometrie beschreibt Beziehungen eines Objektpunkts P zu seinen Bildpunkten P', P'' in zwei perspektiven Bildern. Die Bildpunkte P' und P'' eines Punktes P liegen auf einander entsprechenden (konjugierten) Epipolarlinien. Diese Linien können bei bekannter relativer Orientierung berechnet werden. Wenn ein Objektpunkt in einem Bild vorliegt, liegt er im anderen Bild auf der entsprechenden Epipolarlinie. Bildpaare können normalisiert werden, d.h. die Projektionszentren bleiben erhalten aber die Drehwinkel ω, ϕ, κ werden so bestimmt, dass die Bildebenen zusammenfallen. Dann sind die Bildzeilen parallel zur Basis. Dies erleichtert die Bildzuordnung sehr, da man die Suche nach konjugierten Punkten auf eine Epipolarlinie reduzieren kann.

Der Zusammenhang der Epipolargeometrie kann ohne Kenntnis von Kameraparametern durch die Fundamentalmatrix F beschrieben werden (siehe Epipolarbedingung in Gleichung 2.3).

$$x'^T F x = 0 \quad (2.3)$$

Zur Bestimmung der Fundamentalmatrix sind mindestens sieben homologen Punkte erforderlich. Wenn die Kamerakalibrierung gegeben ist, kann die Epipolargeometrie durch die Elementarmatrix E beschrieben werden (siehe Berechnung von E in Gleichung 2.4 sowie Epipolarbedingung in Gleichung 2.5), was fünf homologe Punkte erfordert (vergleiche [Nister, 2003, 2004]).

$$E = K'^T F K \quad (2.4)$$

$$\hat{x}'^T E \hat{x} = 0 \quad (2.5)$$

Alle notwendigen photogrammetrischen Verarbeitungsschritte sind selbsterklärend, darauf wird in dieser Arbeit nicht näher eingegangen. Die relative Orientierung, also die Positionen der Aufnahmen relativ zueinander im Raum, und die Positionen der identischen Punkte, sowie die Parameter der inneren Orientierung der Kamera, können mit Techniken wie dem *Fünf Punkte Algorithmus* [Nister, 2003, 2004], einer *Bündelblockausgleichung* [Lourakis & Argyros, 2009], *Structure from Motion* [Snavely et al., 2006; Agarwal et al., 2009; Frahm et al., 2010] oder vergleichbaren zielführenden Lösungsansätzen [Strecha et al., 2008; Bartelsen & Mayer, 2010; Deseilligny & Clery, 2011; Bartelsen et al., 2012; Reich et al., 2013; Heinly et al., 2014] bestimmt werden. Im Anschluss daran können aus den Bildern sehr dichte Punktwolken mit Hilfe von unterschiedlichen Matchingmethoden errechnet werden [Hirschmuller, 2008; Bodenmueller, 2009; Bulatov, 2011; Mayer et al., 2012] oder andere Geometrien extrahiert werden [Ulrich, 2003].

2.1.2 Aktuelle Konzepte und Algorithmen für Punktzunordnungen

Der Zusammenhang zwischen zwei sich überlappenden Aufnahmen sind die Punkt-korrespondenzen oder homologen Punkte, welche dieselben abgebildeten Objekte beschreiben. Um zu verstehen, was identische Objektpunkte ausmacht und wie diese bisher gefunden wurden, werden im Folgenden die wichtigsten Grundlagen der verwendeten stereoskopischen Verfahren beschrieben. Anschließend wird dargelegt, warum für eine erfolgreiche Suche nach homologen Punkten insbesondere jene Eigenschaften des Bildes berücksichtigt werden müssen, die projektiv invariant sind.

Um identische Punkte ohne weiteres Vorwissen in zwei Bildern zuordnen zu können, werden Punkte anhand von Kenngrößen beschrieben, die voraussichtlich in beiden Bildern wiederzufinden sind. Ziel ist die Suche nach markanten Punkten, die eindeutig zugeordnet werden können.

Kriterien für solche Merkmale sind nach [Tuytelaars & Mikolajczyk, 2008] und [Luhmann, 2010]:

- Die Deutlichkeit, also die eindeutige Unterscheidbarkeit gegenüber der Nachbarschaft.
- Die Invarianz gegenüber geometrischen und radiometrischen Verzerrungen.
- Die Robustheit, also die Unempfindlichkeit gegenüber Rauschen.
- Die Seltenheit, also eine Unterscheidbarkeit gegenüber anderen Kandidaten.
- Die Interpretierbarkeit, also dass die Merkmale für eine weitere Bildanalyse geeignet sind.

Das Kriterium der geometrischen Invarianz wird im Folgenden genauer betrachtet. Im Bereich der Mathematik ist dieses Thema nicht neu [Patterson, 1928]. Im Bereich der Photogrammetrie werden geometrische Invarianzen sowohl unbewusst, als auch bewusst verwendet. Gängige Merkmalsdeskriptoren wie Harris, SIFT, SURF, FAST, MSER, ORB, BRISK oder ASIFT sind invariant bezüglich Rotationen, Skalierungen oder auch affinen Transformationen. Dies ist ausreichend, solange sich der Blickwinkel gegenüber dem Objekt nicht zu sehr ändert, eine große Kamerakonstante verwendet wird und sich das Objekt nahe der Bildmitte befindet. Unter diesen Bedingungen kann ein stark vereinfachtes affines Kameramodell angenommen werden. Einer der leistungsfähigsten aktuell benutzten Algorithmen ist der SIFT Deskriptor. Bei Veränderungen des Blickwinkels gegenüber dem Objekt sinkt dessen Reproduzierbarkeit im zweiten Bild jedoch schon ab Winkeln von mehr als 20 Grad deutlich ab [Lowe, 1999, 2004]. Diese Abhängigkeit des Merkmals vom annähernd gleichen Blickwinkel ist der große Nachteil dieses Verfahrens. Auch die anderen auf lokale Merkmalspunk-

te bezogenen Deskriptoren und Matchingverfahren verhalten sich ähnlich. Modifizierte Varianten wie beispielsweise Affine-SIFT [Yu & Morel, 2009], die explizit auf größere Veränderungen des Blickwinkels ausgelegt sind, haben als Grundlage ein affines Kameramodell. Damit lassen sich zwar deutlich größere Blickwinkelveränderungen korrekt erkennen, allerdings sind diese auf rein affine Transformationen beschränkt. So können beispielsweise unter der Voraussetzung, dass der Abstand zwischen Kamera und dem Objekt groß ist, Objekte gut beschrieben und wiedererkannt werden. Jedoch ist das Modell der affinen Kamera ein theoretisches Modell, welches nie unter realen Bedingungen zutrifft. Neben den aufgeführten bekannten Algorithmen gibt es noch eine Reihe weiterer Arbeiten, die sich mit der rotationsinvarianten, skaleninvarianten oder der affin-invarianten Merkmalszuordnung beschäftigt haben: [Kliot & Rivlin, 1998; Tao & Grosky, 1998; Mikolajczyk & Schmid, 2001; Araabi *et al.*, 2002; Zhang & Lu, 2002; Sivic & Zisserman, 2003; Tuytelaars & Van Gool, 2004; Rosten & Drummond, 2005; Lin, 2006; Allaire *et al.*, 2008; Wang *et al.*, 2009; Bronstein *et al.*, 2011; Kottman, 2011; Alahi *et al.*, 2012; Bansal *et al.*, 2012; Moisan *et al.*, 2012; Kottman, 2013; Prasad, 2013; Ren *et al.*, 2014; Bryner *et al.*, 2014; Jia *et al.*, 2014; McManus *et al.*, 2014; Verhagen *et al.*, 2014].

2.1.3 Einschränkungen aktueller Konzepte und deren Folgen

Eine sofort sichtbare Folge der Einschränkungen derzeitiger Merkmalsdeskriptoren und Matchingverfahren ist, dass schon bei der Modellierung von scheinbar einfachen Szenen mittels *Structure from Motion* eine Vielzahl an Bildern aufgenommen werden muss, wie etwa bei der Modellierung eines auf dem Boden stehenden rechteckigen Schuhkartons (siehe Skizze in Abbildung 2.1). Um eine erfolgreiche Punktzuordnung durchführen zu können, darf der räumliche Winkel zwischen zwei Bildstrahlen, welche zu identischen Merkmalen in beiden Bildern führen, beispielsweise bei SIFT-Features nicht größer sein als ca. 20 Grad [Lowe, 1999, 2004].

Im Gegensatz zur manuellen photogrammetrischen Auswertung müssen für die automatische Orientierung der Bilder nicht nur die Bilder aufgenommen werden, die ausreichend wären, um jeden Punkt des Schuhkartons in mindestens einem Stereopaar zu erfassen, sondern zusätzlich sehr viele weitere Bilder, die zwischen diesen beiden liegen. Diese Notwendigkeit ergibt sich aus den Bildinformationen, die den Merkmalsbeschreibungen zugrunde liegen. Sie sind bei projektiven Transformationen mit stark unterschiedlichen Blickwinkeln so stark verändert, dass die Merkmale nicht mehr dieselben Kenngrößen aufweisen. Sie können daher nicht mehr eindeutig einander zugeordnet werden. Die beiden Bildbeispiele in Abbildung 2.2 verdeutlichen diese Schwierigkeiten, mit denen punkt- oder blobbezogene Merkmalsdeskriptoren Schwierigkeiten haben. Die lokale Umgebung ändert sich zu stark, als dass identische auf die Textur bezogene Merkmale gefunden werden können.

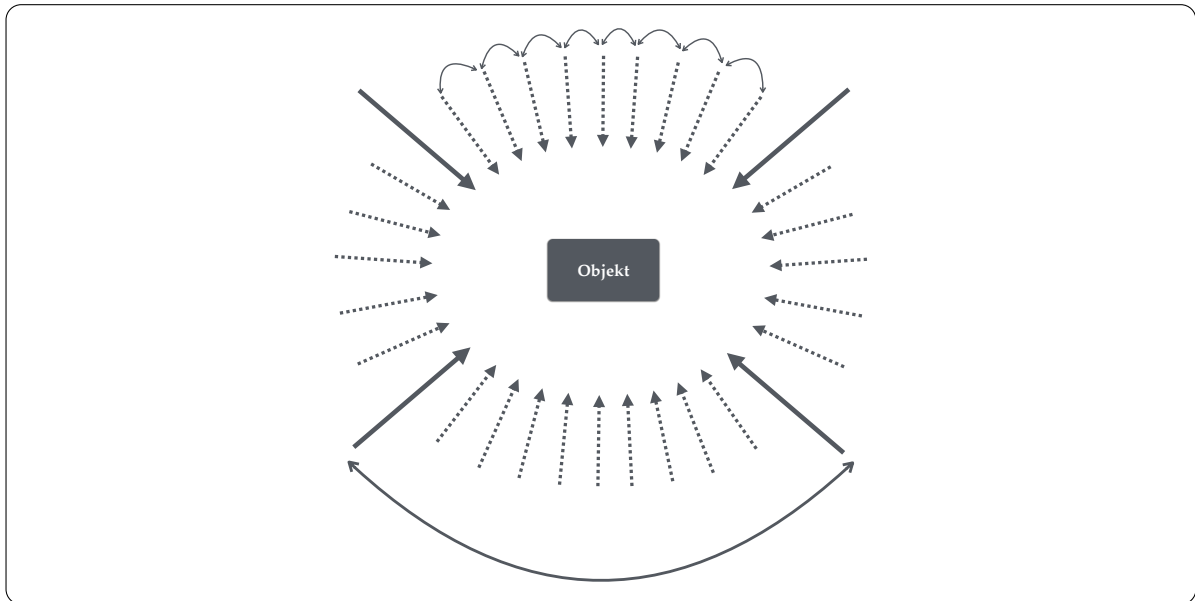


Abbildung 2.1: Grundriss einer Aufnahmesituation eines einfachen Objektes. Bei einer automatischen Auswertung mit aktuellen Techniken müssen außer den vier Aufnahmerichtungen (durchgezogene Linien), welche notwendig wären, um jeden Punkt des Objektes in mindestens zwei Bildern abzudecken, zusätzlich viele weitere Zwischenschritte (gestrichelte Linien) aufgenommen und ausgewertet werden, weil für aktuelle Algorithmen der Winkel zwischen den vier Aufnahmen zu groß ist, da die perspektivische Transformation des Teilbildes nicht berücksichtigt wird.

Daher ist ein Algorithmus, der im Fall von stark konvergenten Aufnahmen überhaupt irgendeine Lösung findet eine echte Neuerung und sehr hilfreich. Mit einem Algorithmus, der identische Punkte bei großen Blickwinkelveränderungen findet, können dann auch bei bestehenden dichten Netzen mit geringen Blickwinkelunterschieden zusätzlich weitere Verknüpfungen für eine Bündelblockausgleichung gefunden werden. So können bestehende Bündelnetze nochmals verbessert werden durch neue Beobachtungen identischer Punkte in Bildern mit stark unterschiedlichem Blickwinkel, was aktuell nicht möglich ist.

Eine weitere Einschränkung vieler Verfahren ist, dass Bilder von deutlich unterschiedlichen Zeitpunkten und mit verändertem Schattenverlauf bei der Suche nach identischen Punkten oftmals keine zufriedenstellenden Lösungen liefern. Hier werden häufig gar keine oder falsche Punktkorrespondenzen gefunden. Auch hochaufgelöste Luftbilder, die aus Flugzeugen mit professionellen Luftbildkameras aufgenommen wurden und deutlich niedriger aufgenommene Bilder, die beispielsweise mit einem UAV (Unmanned Aerial Vehicle) aufgenommen wurden, sind oftmals perspektivisch und radiometrisch so unterschiedlich, dass eine automatische Suche nach identischen Punkten nicht erfolgreich durchführbar ist.

2 Einbettung in den wissenschaftlichen Kontext



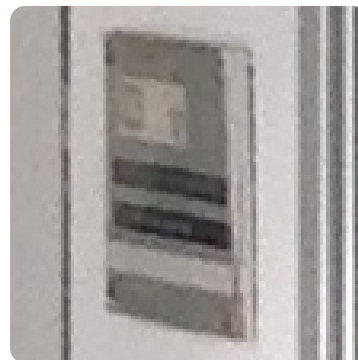
(a) Blickwinkel 1



(b) Blickwinkel 2



(c) Ausschnitt aus (a)



(d) Ausschnitt aus (b)

Abbildung 2.2: Unterschiedliche Blickwinkel verändern die Textur in beiden Bildern, siehe (c) und (d). Obwohl dasselbe Objekt abgebildet wird, können Algorithmen, welche identische Punkte über kleinräumige Features suchen, keine richtigen Ergebnisse liefern.

Folgende Beispiele verdeutlichen die Einschränkungen, die vorhanden sind, wenn aktuell verfügbare Algorithmen für die Suche nach Punktkorrespondenzen benutzt werden. Bei der Aufnahme (Abbildung 2.3) einer Kirche in Rio de Contas in Bahia (Brasilien) waren sehr viele Aufnahmen notwendig, um die relative Orientierung der Bilder zueinander zu bestimmen.

Um das Gebäude photogrammetrisch zu erfassen, wurde es einmal mit einer Kamera im Abstand von ungefähr fünf Metern umrundet. Bei der Umrundung wurde circa nach jedem abgeschrittenen Meter angehalten, um ein oder mehrere Bilder von diesem Standort aus von dem Gebäude aufzunehmen. Für die Suche nach identischen Punkten wurde der SIFT-Algorithmus verwendet, welcher keine starken Blickwinkeländerungen gegenüber dem Objekt zulässt. Wie Lowe in den Veröffentlichungen zum SIFT-Algorithmus bereits beschrieben hat ([Lowe, 1999] S.6 und [Lowe, 2004]

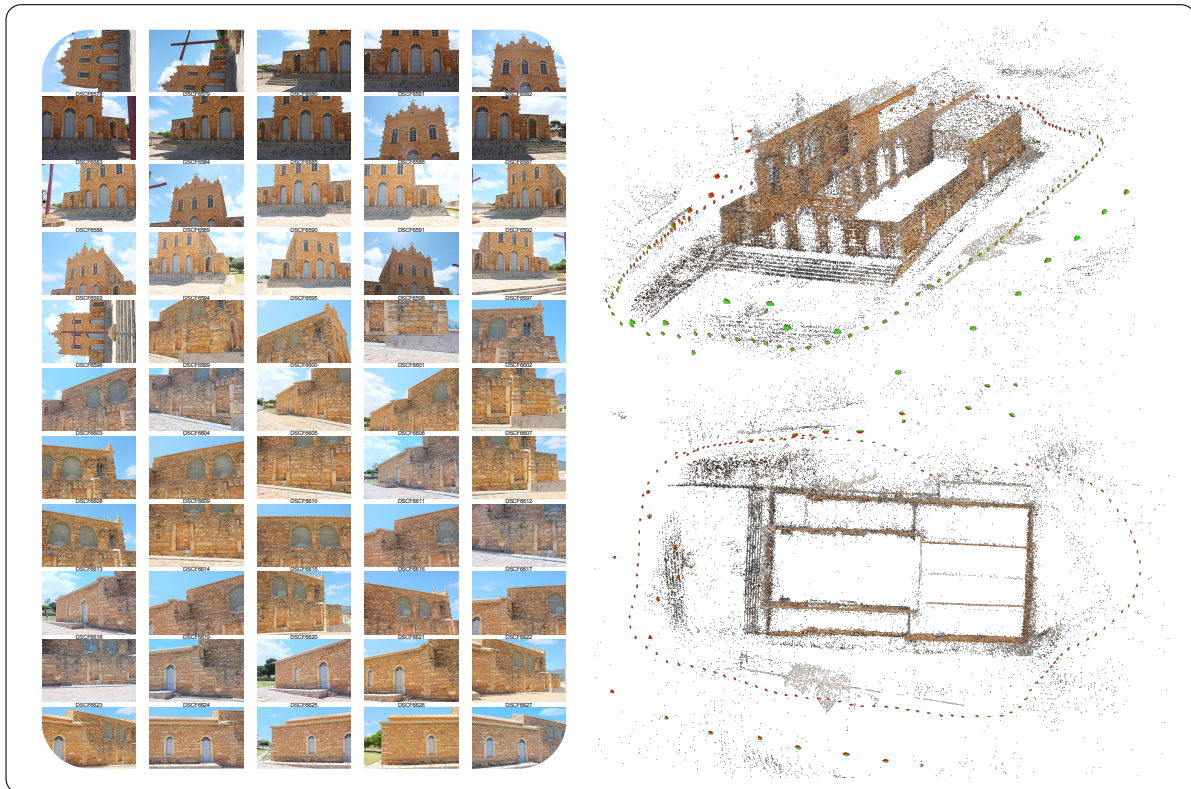


Abbildung 2.3: Die Bilder der Kirche in Rio de Contas in Brasilien (auszugsweise links dargestellt) wurden mit einem sehr hohen Überlappungsbereich und mit jeweils minimaler Änderung des Blickwinkels gegenüber dem Objekt aufgenommen. Für eine erfolgreiche automatische relative Orientierung der Bilder zueinander waren sehr viele Bilder von zahlreichen Standpunkten notwendig. Die errechneten Orientierungen der Bilder sind in den Bildern auf der rechten Seite in rot und grün visualisiert. Die Menge an Standpunkten ist notwendig, um projektive Transformationen im Bild möglichst klein zu halten und eine korrekte Punktzuordnung mit Hilfe des SIFT-Algorithmus durchführen zu können.

S.18 Figure 9) sinkt die Wiederholgenauigkeit bei Winkeln über 20 Grad deutlich ab. Würden alternativ bei der Umrundung des Gebäudes lediglich alle 10 Meter eine oder mehrere Aufnahmen in Richtung des Gebäudes aufgenommen, so können meisten die Punktzuordnungen nicht automatisch mit Hilfe des SIFT-Algorithmus gefunden werden, da hier Blickwinkeländerungen entstehen, die deutlich größer als 20 Grad gegenüber dem Objekt sind. Die zugrunde liegende Bildinformation verändert sich durch die unterschiedliche Perspektive zu stark, um von den kleinräumigen lokalen Merkmalen korrekt zugeordnet werden zu können.

Das zweite Beispiel zeigt eine Fassade eines Fachwerkhouses in Frankfurt am Main (Abbildung 2.4). Diese wurde aus zwei stark unterschiedlichen Richtungen aufge-

nommen. Auch hier findet der SIFT-Algorithmus keine korrekten Zuordnungen, da die deutliche Änderung des Blickwinkels und die dadurch entstehende projektive Transformation eines Teilbereichs des Bildes eine fehlerfreie Suche nach Punktkorrespondenzen verhindern.



(a) Seitliche Aufnahme des gleichen Fachwerkhouses.



(b) Frontalaufnahme eines Fachwerkhouses in Frankfurt am Main.

Abbildung 2.4: Ein Bildbeispiel für einen stark veränderten Blickwinkel zwischen zwei Kameraaufnahmen desselben Objektes. Mit lokalen Merkmalsdeskriptoren wie beispielsweise SIFT können keine korrekten Punktzuordnungen gefunden werden, da sich die kleinräumigen Kenngrößen nach einer projektiven Transformation stark unterscheiden.

Im Folgenden werden die invarianten Eigenschaften vorgestellt, die für eine erfolgreiche Lösung des Problems notwendig sind.

2.2 Projektive Invarianzen

2.2.1 Invariante Kenngrößen bei unterschiedlichen Transformationen

Um invariant gegenüber geometrischen Transformationen zu sein, müssen die Merkmalsbeschreibungen invariant gegenüber der jeweiligen Transformation sein, die notwendig ist, um den Teilbereich des Bildes in den entsprechenden Teilbereich des anderen Bildes zu transformieren. Beispielsweise können folgende Kenngrößen im Bild bestimmt werden und als Bildbeschreibung dienen: Richtungen, Entfernungen, Richtungsunterschiede, Verhältnisse von Entfernungen, Parallelitäten, Längenverhältnisse von parallelen Liniensegmenten, geometrische Schwerpunkte, Inzidenzen, gerade Linien, Konnektivitäten und Doppelverhältnisse. In der Hierarchie der geometrischen Transformationen (Abbildung 2.5) gehen mit jeder weiterführenden Transfor-

mation mehrere invariante Kenngrößen verloren. Um dies zu verdeutlichen, sind im Folgenden einige Bildbeispiele der Transformationen und den enthaltenen invarianten Eigenschaften aufgeführt (Abbildungen 2.6, 2.7, 2.8 und 2.9).

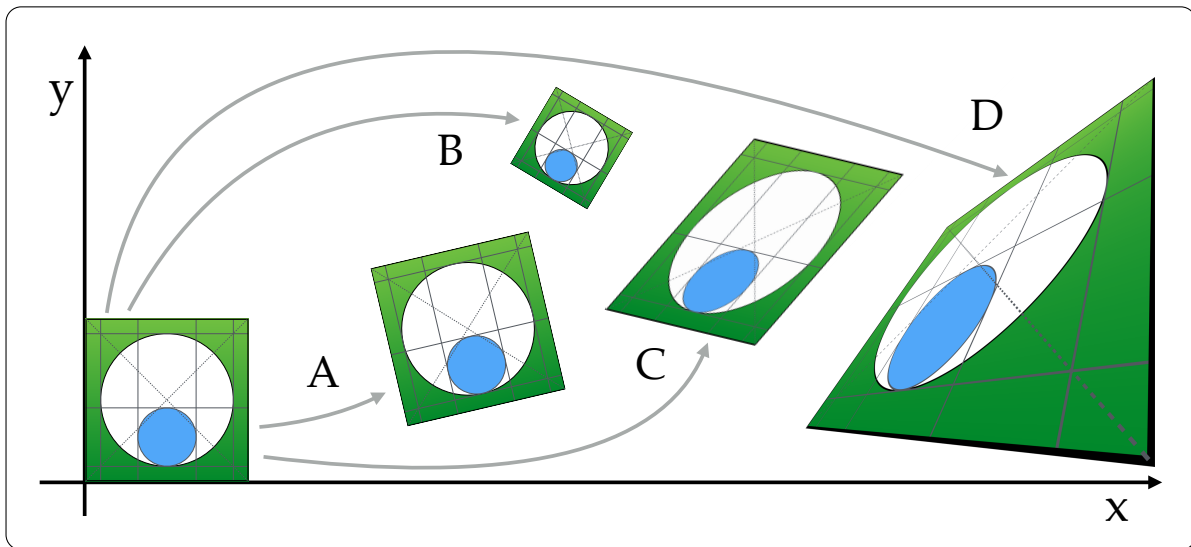


Abbildung 2.5: Eine schematische Darstellung der Transformationen der jeweiligen Geometrien. A: Euklidische Transformationen, B: Ähnlichkeitstransformation, C: Affine Transformation, D: Projektive Transformation.

Die euklidische Transformation

Bei einer Euklidischen Transformation (Abbildung 2.6), also einer Translation und einer Rotation, bleiben folgende invariante Eigenschaften erhalten: *Entfernungen, Winkel, Verhältnisse von Entfernungen, Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten*. Neben den invarianten Eigenschaften, die auch bei den anderen Transformationen vorhanden sind, sind hier zusätzlich Entfernungen direkt vergleichbar. Die euklidischen Distanzen zwischen zwei identischen Punkten sind sowohl im Ursprungsbild, als auch im transformierten Bild gleich groß.

Die Ähnlichkeitstransformation

Die Ähnlichkeitstransformation (Abbildung 2.7), die eine Translation, eine Rotation und eine isotrope Skalierung abdeckt, hat folgende invariante Eigenschaften: *Winkel, Verhältnisse von Entfernungen, Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten*. Im Gegensatz zur euklidischen Transformation sind hier keine Strecken mehr vergleichbar. Diese können durch die

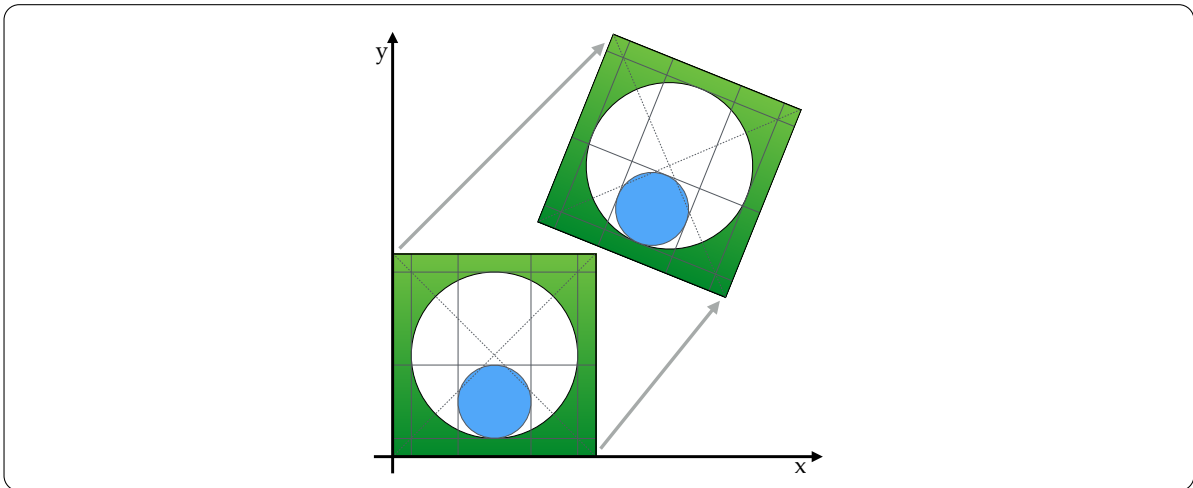


Abbildung 2.6: Die euklidische Transformation bestehend aus Translation und Rotation. Die invarianten Eigenschaften der euklidischen Transformation sind: Entfernungen, Winkel, Verhältnisse von Entfernungen, Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.

Skalierung unterschiedlich groß sein. Wie schon bei der euklidischen Transformation sind bei der Ähnlichkeitstransformation die Winkel im Bild gleich groß. Das bedeutet beispielsweise, dass die Differenzen der Richtungswinkel, die zwischen mehreren identischen Punkten im Ausgangsbild und im transformierten Bild gemessen wurden, identisch sind und sich direkt vergleichen lassen.

Die affine Transformation

Die affine Transformation, bestehend aus einer Translation, einer Rotation, einem Skalierungsfaktor je Koordinatenachse und einer Scherung, bewahrt folgende invariante Eigenschaften: *Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten*. Einer der Vorteile der affinen Geometrie ist, dass Parallelitäten und geometrische Schwerpunkte invariant gegenüber Transformationen sind. Vor allem der geometrische Schwerpunkt ist Bestandteil vieler Methoden, wie beispielsweise der Delaunay Triangulation [Delaunay, 1934]. In einer affinen Geometrie bleibt dieser auch nach einer Transformation erhalten.

Die projektive Transformation

Bei der projektiven Transformation bleiben lediglich *Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten* als Invariante erhalten. Parallele Linien gehen verloren

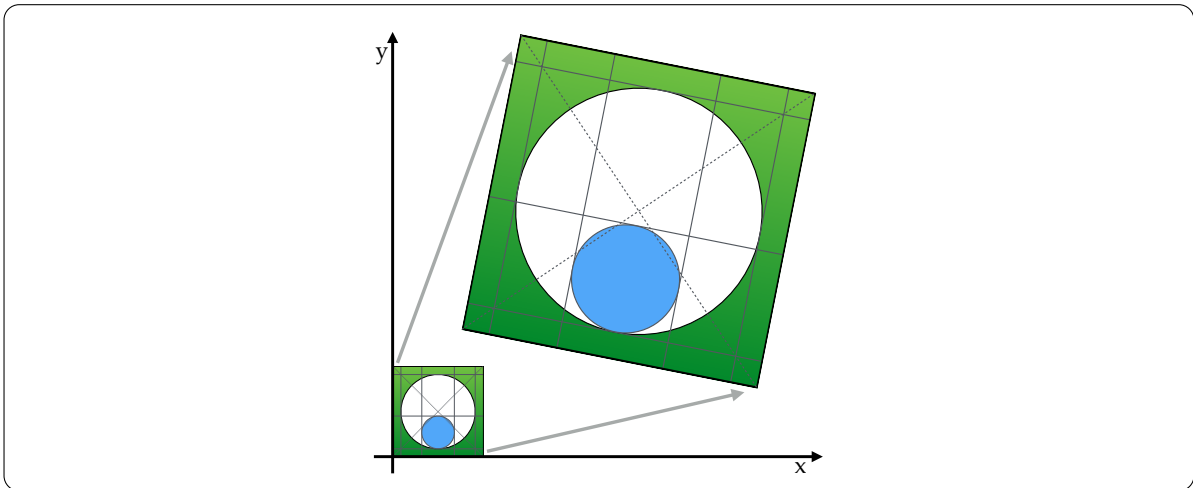


Abbildung 2.7: Die Ähnlichkeitstransformation bestehend aus Translation, Rotation und einer isotropen Skalierung. Die invarianten Eigenschaften der Ähnlichkeitstransformation sind: Winkel, Verhältnisse von Entfernungen, Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.

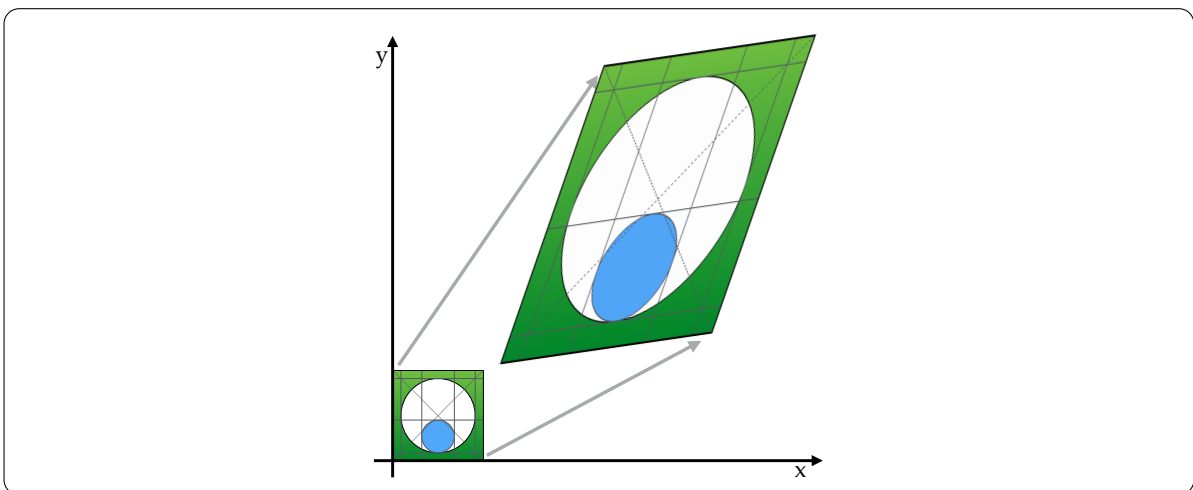


Abbildung 2.8: Die affine Transformation bestehend aus Translation, Rotation, Skalierungen entlang der beiden Achsen und einer Scherung. Die invarianten Eigenschaften der affinen Transformation sind: Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.

und geometrische Schwerpunkte sind nicht konstant. Eine Folge daraus ist, dass jene Algorithmen, die auf geometrischen Schwerpunkten aufbauen, bei starken Veränderungen des Blickwinkels und bei projektiv abgebildeten Bildern fehlerhafte oder

keine Ergebnisse liefern.

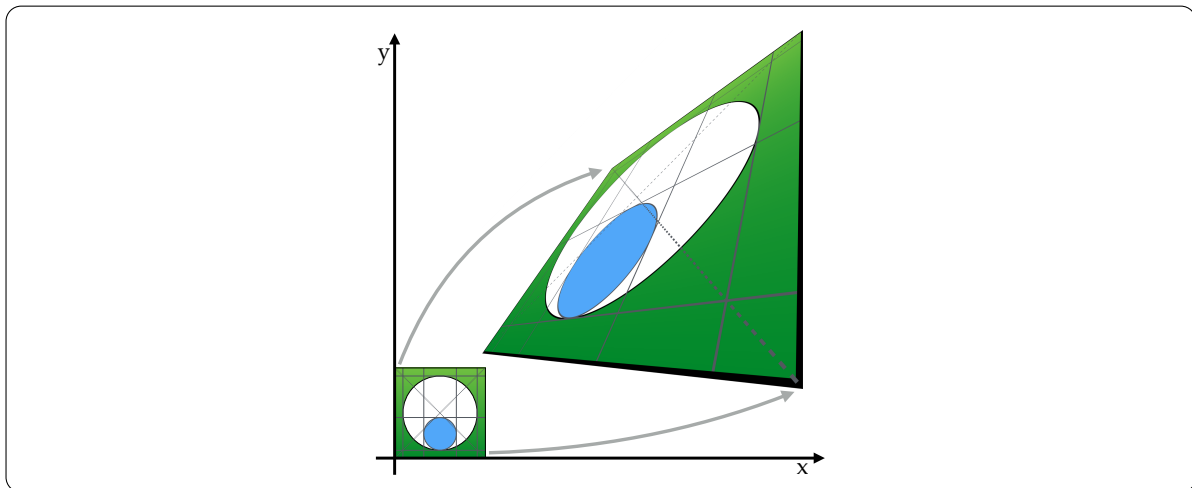


Abbildung 2.9: Die projektive Transformation, bestehend aus Translation, Rotation, Skalierungen entlang der beiden Achsen, einer Scherung und der perspektivischen Projektion. Die invarianten Eigenschaften der affinen Transformation sind: Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.

In Tabelle 2.1 sind die Invarianzen der jeweiligen Transformation zusammengefasst. Daraus ergeben sich unmittelbar einsichtig nach [Mikhail *et al.*, 2004; Faugeras & Luong, 2004] folgende Aussagen. Bei projektiven Transformationen bleiben nur Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten als Invariante übrig. Alle anderen Kenngrößen entfallen. Die invarianten Eigenschaften der projektiven Transformation sind auch invariant bei den darin enthaltenen Transformationen der Euklidischen-, der Ähnlichkeits- und der affinen Geometrie. Eine Merkmalsbeschreibung, die ausschließlich projektiv invariante Merkmale benutzt, ist für alle aufgeführten Transformationen und Geometrien gültig.

2.2.2 Aktuelle Methoden basierend auf projektiven Invarianzen

Die Idee, projektiv invariante Kenngrößen zur Bildbeschreibung zu benutzen, wurde bereits mehrfach vorgeschlagen und mögliche Lösungswege wurden erarbeitet [Weiss, 1988; Forsyth *et al.*, 1990; Barrett *et al.*, 1991; Forsyth *et al.*, 1991; Coelho *et al.*, 1992; Mundy & Zisserman, 1992; Weiss, 1993; Gevers & Smeulders, 1994; Hartley, 1994; Carlsson, 1998; Goodall & Mardia, 1999; Lourakis *et al.*, 2000; Li *et al.*, 2002; Li & Lee, 2002; van Liere & Mulder, 2003; Wang *et al.*, 2004; Zhang & Lu, 2004; Lu & Manduchi, 2004; Koser & Koch, 2007]. Die auf diesem Gebiet wichtigsten Methoden sind im Folgenden detailliert betrachtet. In [Forsyth *et al.*, 1990] wurde zusätzlich zu den in Tabelle 2.1 aufgeführten Invarianten Kenngrößen folgende projektiv

	Euklidische Geometrie	Ähnlichkeits-Geometrie	Affine Geometrie	Projektive Geometrie
Verwendete Transformationen:				
Rotation	✓	✓	✓	✓
Translation	✓	✓	✓	✓
Isotrope Skalierung		✓	✓	✓
Skalierung entlang der Achsen			✓	✓
Scherung			✓	✓
Perspektivische Projektionen				✓
Invariante Eigenschaften:				
Entfernungen	✓			
Winkel	✓	✓		
Verhältnisse von Entfernungen	✓	✓		
Parallelitäten	✓	✓	✓	
Geometrische Schwerpunkte	✓	✓	✓	
Inzidenzen	✓	✓	✓	✓
Doppelverhältnisse	✓	✓	✓	✓
Gerade Linien	✓	✓	✓	✓
Konnektivität	✓	✓	✓	✓

Tabelle 2.1: Eine Übersicht über die unterschiedlichen Typen von Geometrien in Abhängigkeit der verwendeten geometrischen Transformationen und der unterschiedlichen invarianten Eigenschaften (nach [Mikhail *et al.*, 2004] sowie [Faugeras & Luong, 2004], verändert). Im oberen Bereich der Tabelle sind die jeweiligen verwendeten Transformationen je Geometrie dargestellt. Im unteren Bereich sind die mit den jeweiligen Geometrien verbundenen invarianten Eigenschaften aufgelistet. Es ist sofort ersichtlich, dass mit jeder hinzukommenden Transformation invariante Eigenschaften verloren gehen. Bei der projektiven Geometrie bleiben lediglich Inzidenzen, Doppelverhältnisse, gerade Linien sowie Konnektivitäten als projektiv invariante Eigenschaften erhalten. Diese Invariante sind auch für alle anderen aufgeführten Geometrien und deren Transformationen gültig.

planare Invariante untersucht: Fünf koplanare Punkte, zwei Linien und zwei Punkte, zwei koplanare Kegelschnitte, und die nicht euklidische Metrik. Insbesondere die Positionsbestimmung von dreidimensionalen Objekten vor der Kamera wurde mit Hilfe der Invarianten der zwei Kegelschnitte untersucht. Diese invariante Kenngröße ist vor allem bei industriellen Anwendungen von Bedeutung, wenn Werkstücke mit kreisförmigen ebenen Öffnungen zuverlässig in Bildern mit stark unterschiedlichem Blickwinkel wiederzufinden sind.

Auch [Barrett *et al.*, 1991] befasste sich mit der grundlegenden Bestimmung von projektiven Invarianten in Bildern. Zusätzlich zu den koplanaren Objektpunkten wurden hier auch Objektpunkte, die im dreidimensionalen Raum liegen und nicht koplanar sind, mit Hilfe von Doppelverhältnissen beschrieben. Dafür sind sechs Punkte im Raum notwendig. Diese sechs Punkte im Raum für die Bestimmung von projektiven Invarianten mehrerer Bilder zu nutzen, ist auch Gegenstand der Untersuchung bei [Quan, 1995]. Er konnte zeigen, dass hierfür mindestens drei Bilder notwendig sind, und dass bei der Berechnung der Invarianten aus sechs Punkten im Raum bis zu drei Lösungen vorkommen.

[Coelho *et al.*, 1992] haben in mehreren Experimenten nachgewiesen, dass es möglich ist, projektiv planare invariante Merkmale anhand ihrer Kenngrößen in projektiv transformierten Bildern zuzuordnen. Wenn dieselben Eckpunkte in zwei Bildern vorgegeben wurden, konnten diese durch verschiedene Arten der projektiv invarianten Merkmale bestätigt werden. Diese Merkmale umfassten das Doppelverhältnis von vier Punkten entlang einer projektiven Geraden, die zwei projektiven Invarianten bei fünf Punkten auf einer projektiven Ebene, die projektive Invariante bei zwei Linien und zwei Punkten und die projektive Invariante bei zwei koplanaren Kegelkurven.

[Meer *et al.*, 1994] haben sich mit der Suche nach Punktkorrespondenzen, welche mit Hilfe von P^2 -Invarianten (siehe auch Kapitel 3.1.3 sowie Kapitel 3.3) gefunden wurden, in ihrer Arbeit auseinandergesetzt. Der vorgestellte Algorithmus war eine *blinde Technik*, die außer den reinen Punktkoordinaten von manuell extrahierten Bildpunkten keine weiteren Informationen berücksichtigt hat. Mit einer *Brute Force Strategie* wurden alle Kombinationsmöglichkeiten der Fünf-Punkte-Kombinationen durchprobiert, um identische Punkte auf einer projektiven Ebene in zwei unterschiedlichen Bildern zuzuordnen. Mit diesem Ansatz können korrekte Punktzuordnungen unter einem extrem hohen Rechenaufwand gefunden werden (siehe Kapitel 3.3.1), wenn in beiden Bildern die gleichen Eckpunkte manuell vorgegeben wurden.

[Meer *et al.*, 1998] zeigten Möglichkeiten der effizienten Indizierung von P^2 -Invarianten in Bildern auf. Mittels geometrischem Hashing wurden P^2 -Invariante, also die zwei Invarianten aus fünf koplanaren Punkten (siehe auch Kapitel 3.1.3 sowie Kapitel 3.3), sowie P^3 -Invariante, also sechs Punkte im Raum, indiziert und für die Suche nach Punktkorrespondenzen eingesetzt.

[Branca *et al.*, 1998] setzt P^2 -Invariante bei der autonomen Navigation von Unterwasserfahrzeugen ein. Er löst die Punktzuordnung mittels Graphen-Matching. Da-

bei werden alle Untermengen von fünf Punkten und deren P^2 -Invarianten in beiden Bildern miteinander verglichen. Die Autoren weisen auf die Wichtigkeit robuster, radiometrisch invarianter Features hin, die unabhängig von der jeweiligen Beleuchtung sind, was bei Einsätzen mit Tauchrobotern, die ausschließlich künstliches Licht zur Verfügung haben, elementar ist. Hierfür wurden Moravec Features [Moravec, 1977] benutzt.

Im Bereich der Wiedererkennung von Konturen einzelner Symbole, wie etwa Buchstaben, welche projektiv verzerrt wurden, sind in den letzten Jahren große Fortschritte erzielt worden. Bereits [Putjarupong *et al.*, 2004] verwendete P^2 -Invariante zur Beschreibung von Konturen. [Li & Tan, 2010] benutzt eine Folge von Spektren von Doppelverhältnissen (*cross ratio spectrum*) an einem einzelnen ebenen Symbol, beispielsweise einem Buchstaben, der perspektivisch transformiert wurde. Dabei werden zunächst markante Punkte, welche die konvexe Hülle der Kontur des Buchstaben vor und nach der Transformation beschreiben, bestimmt. Von jedem Eckpunkt wird zu den anderen Eckpunkten entlang einer Linie der Helligkeitsverlauf ermittelt und lokale Maxima in deren Verläufen werden bestimmt. Um ein Doppelverhältnis ausrechnen zu können, werden vier Punkte entlang einer projektiven Geraden benutzt (siehe auch Kapitel 3.1.3). Der beschriebene Algorithmus benutzt für die vier Punkte des Doppelverhältnisses den Anfangs- und Endpunkt der Linie zwischen den beiden Punkten, welche Teil der Konturbeschreibung sind, und jeweils die ersten beiden Maxima in den Helligkeitsverläufen derselben Verbindungslinie. Die Zuordnung der identischen Punkte in beiden Bildern erfolgt über eindeutige identische Doppelverhältnisse. Dadurch, dass nur die ersten beiden Maxima auf der Verbindungslinie zwischen zwei Punkten zur Beschreibung der Kontur benutzt werden, eignet sich dieses Verfahren lediglich für Objekte mit geringer Komplexität.

[Luo *et al.*, 2013] führt den Begriff der *Charakteristischen Zahl* (*Characteristic Number*) ein, die als Generalisierung des Doppelverhältnisses angesehen werden kann. Mit der Charakteristischen Zahl können Konturen innerhalb einer abgeschlossenen konvexen Hülle projektiv invariant beschrieben werden. Diese kann beim Wiedererkennen projektiv transformierter abgeschlossener Symbole verwendet werden. Hierzu wird eine konvexe Hülle um die zu untersuchende Form gelegt. Von allen Punkten auf der konvexen Hülle werden Linien zu allen anderen Punkten auf der Hülle erstellt und entlang deren Verlauf wird nach Kanten im Gradientenbild gesucht. Im Gegensatz zum Ansatz der Spektren von Doppelverhältnissen werden hier alle Schnittpunkte des Gradientenbildes auf den Linien mit berücksichtigt. In [Fan *et al.*, 2014] wurde dann gezeigt, dass diese Charakteristische Zahl projektiv invariant ist. Die Charakteristische Zahl findet bei [Fan *et al.*, 2015] Anwendung in der Erstellung von Gesichtsmarkmalen, die dazu dient, Gesichter in Bildern zu lokalisieren und diese dann zuzuordnen.

2.3 Radiometrische Invarianzen

Der Hauptteil dieser Arbeit liegt in der Geometrie und deren Invarianzen, was ein Teilbereich der Mathematik ist. Der Themenbereich der Radiometrie ist eigentlich ein komplett anderer, nämlich ein Teil der Physik, der sich mit der elektromagnetischen Strahlung befasst. Diese ist nicht Teil dieser Arbeit. Die abgebildete Geometrie innerhalb eines Bildes würde jedoch ohne die Radiometrie gar nicht existieren. Durch die Tatsache, dass die verwendeten Bilder durch ein digitales Aufnahmesystem entstanden sind, dessen Eingangsdaten elektromagnetische Strahlung ist, steht die abgebildete Geometrie in direkter Beziehung zur Radiometrie. Nach [Luhmann, 2010] (S.165) verarbeitet ein digitales Aufnahmesystem elektromagnetische Strahlung mit Hilfe eines Bildsensors und der Signalverarbeitung zu einem analogen Bildsignal, welches anschließend digital abgetastet und quantisiert wird. Das Ergebnis ist ein digitales Bildsignal, welches als Bilddatei gespeichert wird. Abhängig von der reflektierten Strahlung, die von einem Objekt zurückgeworfen wird, verändert sich auf diese Weise die radiometrische Information der abgetasteten Bildwerte in den entsprechenden Bildkanälen. Sowohl die unterschiedliche Reflektanz des Objektes, als auch die darauf treffenden unterschiedlichen Strahlungsintensitäten haben auf diese Weise einen direkten Einfluss auf die Abbildung der Geometrie im Bild. In diesem Kapitel werden Möglichkeiten aufgezeigt, wie trotz der großen Veränderungen einzelner Intensitätswerte, welche dieselben Objekte in unterschiedlichen Bildern beschreiben, Radiometrische Invarianzen gefunden und verwendet werden können.

2.3.1 Herausforderungen radiometrisch veränderter Bildinformation

Radiometrisch invariante Merkmale sind zum einen notwendig, um zuverlässig identische Punkte in Bildern mit unterschiedlicher Beleuchtung zu finden. Zum anderen sind radiometrisch invariante Merkmale bei Bildern, die mit unterschiedlichen Sensoren aufgenommen wurden, besonders wichtig.

Bei der ersten Gruppe (Bildern mit unterschiedlicher Beleuchtung), beeinflussen mehrere äußere Faktoren die Radiometrie. Dies können künstliche Beleuchtungsquellen sein, welche sich räumlich und zeitlich verändern können. Zwei Bilder, die im gleichen Innenraum aufgenommen wurden, sehen beispielsweise anders aus, wenn zwischen den Aufnahmen eine künstliche Lichtquelle an- beziehungsweise ausgeschaltet wurde. Auch Bilder, bei denen ein Blitz benutzt wurde, haben unterschiedliche radiometrische Eigenschaften im Gegensatz zu solchen ohne Blitz. Weitere Einflussfaktoren sind Wetter und Sonnenstand. Allein die Tatsache, ob unmittelbar Sonnenlicht auf eine Oberfläche scheint, oder ob dieselbe Oberfläche diffuses Licht bei bewölktem Himmel reflektiert, wirkt sich auf die Bildinformation aus. Bei Sonnenlicht können

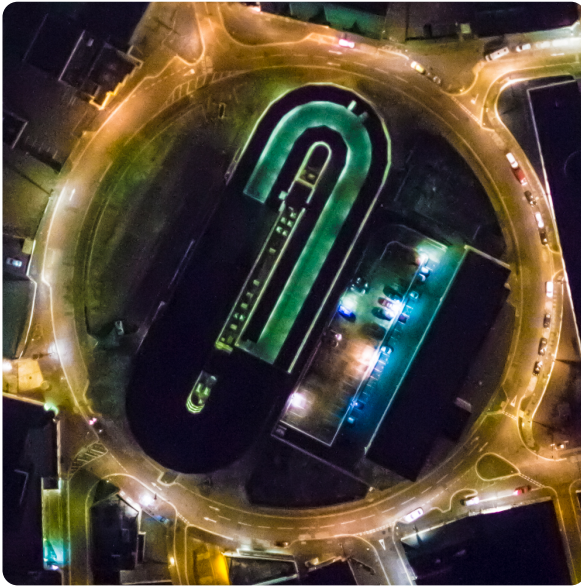
unterschiedliche Zeitpunkte am Tag und während des Jahres zu völlig unterschiedlichen Schattenwürfen von Objekten im Bild führen. Der Schatten wirkt sich dabei in zweierlei Hinsicht aus: zum einen wird die verschattete Stelle dunkler und zum anderen entsteht ein neuer Gradient im Bild an der Stelle, wo der Schatten auf der Oberfläche endet und daneben Sonnenlicht auf die Oberfläche trifft. Am deutlichsten ist der Unterschied sichtbar bei Tag- und Nachtaufnahmen des gleichen Gebietes, wie es in Abbildung 2.10 zu sehen ist. Hier ist die Veränderung besonders stark. Sonnenlicht wird durch künstliches Licht ersetzt und Schatten und helle beleuchtete Bereiche variieren deutlich. Auch die Intensitäten der einzelnen Farbkanäle ändern sich durch die unterschiedlichen Lichtquellen und die reflektierenden Eigenschaften der Oberflächen. Was in einem Bild grau erscheint, wie zum Beispiel das Dach des Gebäudes in der Bildmitte, ist im anderen Bild grünlich angestrahlt.

Zu der zweiten Gruppe (Bilder, die aus unterschiedlichen Sensoren stammen) gehören beispielsweise Bilder, die mit den drei Farbkanälen Rot, Grün und Blau (RGB) aufgenommen wurden. Die meisten Consumer Kameras produzieren diese RGB-Bilder. Daneben gibt es noch viele weitere Bilder wie etwa Infrarotbilder, Wärmebilder oder 2.5-D Tiefenbilder aus Time-of-flight Kameras (z.B. MESA SR4000 oder PMD Cam-Cube). Die unterschiedlichen Bildtypen enthalten jeweils andere Informationen der abgebildeten Objekte einer Szene. Ein allgemeines Verfahren, das eine Punktzuordnung zwischen Bildern ermöglicht, die mit unterschiedlichen Sensoren aufgenommen wurden, kann helfen, auch automatisch Punktkorrespondenzen in diesen Bildern zu finden.

2.3.2 Aktuelle Methoden basierend auf radiometrischen Invarianzen

Neue Merkmalsdeskriptoren zu entwerfen, die invariant gegenüber den in Kapitel 2.3.1 angesprochenen radiometrischen Veränderungen sein sollen, ist in den letzten Jahren vor allem mit dem Ansatz verfolgt worden, monotone Intensitätsänderungen zu berücksichtigen.

Beispielsweise beschreibt [Gupta & Mittal, 2008] einen Merkmalsdeskriptor, der gegenüber monotonen Änderungen der Intensitäten invariant ist. In [Tang *et al.*, 2009] wird ein Merkmalsdeskriptor vorgestellt, der lokale Patches benutzt und auch invariant gegenüber monotonen Intensitätsänderungen ist. [Gupta *et al.*, 2010] widmet sich diesbezüglich vor allem Änderungen, die bei gaußschem Rauschen auftreten. Die Merkmale bei [Wang *et al.*, 2011] sind invariant gegenüber Bildrotationen und monotonen Intensitätsänderungen. Der neueste Ansatz von [Verdie *et al.*, 2015] beschreibt ein lernbasiertes Verfahren, mit dem identische Punkte in Tag- und Nachtbildern gefunden werden können. Das Verfahren benötigt jedoch eine Vielzahl an Bildern von exakt der gleichen Aufnahmeposition zum Anlernen der Merkmale, was sich vor allem bei statischen Webcam Zeitreihen praktikabel verwirklichen lässt.



(a) Radiometrie bei Nacht.



(b) Radiometrie bei Tag.

Abbildung 2.10: Beispiele für unterschiedliche radiometrische Bildinformation bei verändertem Aufnahmezeitpunkt. Es ändern sich unter anderem: die Intensitäten der Farbkanäle, Schattenverläufe und Helligkeiten. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)

2.4 Abgeleitete Schlüsse für den neuen Bildzuordnungs-Ansatz

2.4.1 Neue Beiträge dieser Arbeit

Der neue Lösungsansatz zum Auffinden von Punktkorrespondenzen leitet sich aus folgenden Schlüssen ab. Die in den Kapiteln 2.1.2 und 2.1.3 erörterten derzeit verwendeten Konzepte und Algorithmen benutzen punkt- oder blobbezogene Merkmalsbeschreibungen zum Vergleichen von Bildinhalten. Diese Methoden funktionieren in ihren Anwendungsgebieten jeweils ausgezeichnet und sie produzieren verlässliche Ergebnisse. Sie sind robust, stabil, haben eine hohe Wiederholgenauigkeit und schnelle Implementierungen. Die Kehrseite dieser Algorithmen ist, dass sie maximal invariant gegenüber affinen Transformationen sind und keine beliebig starken Veränderungen des Blickwinkels gegenüber einem Objekt zulassen. Zudem sind sie anfällig gegenüber radiometrischen Veränderungen im Bild.

Die generelle Idee für den Lösungsansatz ist daher folgende: Neue vergleichbare Bildmerkmale müssen invariant gegenüber der projektiven Transformation sein, die notwendig ist, den zugehörigen Teilbereich des ersten Bildes in den des zweiten Bildes zu transformieren. Zur Lösung dieser Aufgabe werden Bildbeschreibungen erstellt, die ausschließlich auf projektiven und radiometrischen Invarianzen beruhen (vergleiche Kapitel 3.1 und 3.2). Diese Bildbeschreibungen werden im Anschluss verglichen und identische Merkmale werden als Punktkorrespondenzen ausgegeben.

2.4.2 Unterschiede zu bestehenden Arbeiten

Die Unterschiede zu bestehenden Arbeiten sind im Folgenden beschrieben. Der Algorithmus für die Suche nach Punktkorrespondenzen benutzt projektiv invariante Kenngrößen, die in Kapitel 2.2.1 besprochen wurden. Bisherige Arbeiten (vergleiche Kapitel 2.2.2) benutzen auch einzelne projektiv invariante Kenngrößen. Die Neuheit dieser Arbeit ist, dass ausschließlich projektive Invarianzen benutzt werden und diese zusätzlich miteinander kombiniert werden. Zudem dienen als Eingangsdaten ausschließlich Merkmale, die in der Mehrzahl als radiometrisch invariant angesehen werden können, welche in Kapitel 2.3 aufgeführt sind. Weiterhin wird ein allgemein übertragbares Konzept erstellt, das in unterschiedlichsten Anwendungsbereichen benutzt werden kann, und das ohne weiteres Vorwissen über eine Szene einsetzbar ist. Der Algorithmus kann auch mit partiellen Verdeckungen oder einzelnen Bildausschnitten umgehen, was bei [Luo *et al.*, 2013] nicht möglich ist und was bei vorherigen Ansätzen nur mit einer extrem hohen Zahl an Rechenoperationen möglich war. Ein weiterer Unterschied ist das neue Konzept der punktbezogenen Zuordnung von virtuellen Linienfeatures, welches Doppelverhältnisse als projektiv invariante Merkmale benutzen. Durch diese neue Zuordnungsmethode kann der Rechenaufwand gegenüber bestehenden projektiv invarianten Ansätzen wie zum Beispiel [Meer *et al.*, 1994] drastisch verringert werden, da hier nicht mehr $\binom{n}{5}$ Rechenoperationen je n Featurepunkte im Bild durchgeführt werden müssen, sondern nur noch $\binom{n}{2}$ (siehe Kapitel 3.3.1 und Kapitel 6.5).

3 Grundlagen und Konzeption

Im folgenden Kapitel werden die notwendigen Grundlagen beschrieben, um Bildmerkmale zu erstellen, die projektiv- und radiometrisch invariant sind. Im ersten Teil werden die Eigenschaften zusammengefasst, die bei projektiven Transformationen invariant sind und im zweiten Teil jene, die radiometrisch invariant sind.

3.1 Konzeption projektiver Invarianzen

Wie in Kapitel 2.2.1 gezeigt worden ist, darf eine allgemein gültige Merkmalsbeschreibung und der anschließende Zuordnungsprozess nur Eigenschaften enthalten, die mit Hilfe der projektiv invarianten Merkmale beschrieben werden. Im Folgenden werden diese Eigenschaften genauer vorgestellt. Jedes dieser Merkmale bezieht sich auf räumliche Zusammenhänge zwischen mehreren diskreten Punkten im Bild.

Konsequenterweise dürfen die Funktionen und Merkmale, welche nicht projektiv invariant sind, wie etwa Entfernungen, Winkel, Verhältnisse von Entfernungen, Parallelitäten oder geometrische Schwerpunkte, nicht bei der Suche nach Punktkorrespondenzen benutzt werden. Dazu zählen auch Trinangulationsalgorithmen wie die Delaunay Triangulation [Delaunay, 1934], da bei dieser die Punkte der gebildeten Dreiecke von einer Umkreisbedingung abhängen, was einer Abhängigkeit vom geometrischen Schwerpunkt entspricht, welcher nicht projektiv invariant ist.

3.1.1 Gerade Linien

Eine Eigenschaft der Zentralprojektion ist die Kollinearität. Aus den Kollinearitätsgleichungen (Gleichungen 2.1 und 2.2) ergibt sich, dass Geraden im dreidimensionalen Raum auch als Geraden im Bildraum abgebildet werden [Luhmann, 2010]. In Abbildung 3.1 ist dies beispielhaft dargestellt. Das dreidimensionale Lineal mit seinen geraden Kanten wird aus zwei unterschiedlichen Richtungen fotografiert. Bei einem verzeichnungsfreien Bild bilden sich die geraden Kanten des Lineals als ebenfalls gerade Kanten im Bild ab. Dies ist unabhängig von der Blickrichtung auf das Lineal. Die Eigenschaft der Kollinearität ist Voraussetzung für alle anderen projektiv invarianten Merkmale, die im Folgenden beschrieben werden.



(a) Schrägansicht eines Lineals.



(b) Aufsicht auf dasselbe Lineal.

Abbildung 3.1: Die gerade Linien der Kanten eines Lineals bleiben gerade Linien bei einer perspektivischen Projektion. Dies ist unabhängig von der Blickrichtung auf das Objekt.

3.1.2 Inzidenzen

Eine geometrische Inzidenz ist eine projektiv invariante Eigenschaft, welche das Auftreten von geometrischen Elementen und deren Beziehungen zueinander beschreibt (siehe auch [Pach & Sharir, 2004]). Ein Punkt P_3 kann beispielsweise als Teilmenge der Punkte einer Strecke $\overline{P_1P_2}$ auftreten (siehe Abbildung 3.2). Ist der Punkt P_3 ein Teil der Strecke $\overline{P_1P_2}$, so ist auch der perspektivisch transformierte Punkt P'_3 Teil der Strecke $\overline{P'_1P'_2}$. Ist ein Punkt nicht Teil der Geraden oder Strecke zwischen zwei Punkten, so ist er auch nach seiner Transformation nicht Teil dieser Geraden oder Strecke.

Gleiches gilt für das Auftreten von Punkten in Ebenen. Ist ein Punkt P Teil einer projektiven Ebene E , so ist auch der projektiv transformierte Punkt P' Teil der gleichermaßen transformierten Ebene E' (siehe Abbildung 3.3). Wenn ein Punkt nicht Teil einer Ebene ist, so ist der Punkt auch nach der projektiven Transformation nicht Teil dieser Ebene. Liegt ein Punkt P innerhalb einer konvexen Hülle, so liegt der transformierte Punkt auch innerhalb der transformierten konvexen Hülle. Umgekehrt bleiben Punkte, die außerhalb einer konvexen Hülle liegen auch nach einer Transformation außerhalb der transformierten konvexen Hülle.

Übertragen auf den Anwendungsfall in der projektiven Geometrie bedeutet dies beispielsweise, dass die Lage eines Punktes dadurch beschrieben werden kann, dass er auf dem Schnittpunkt von zwei Strecken oder Geraden liegt. Ein Punkt P_S der auf dem Schnittpunkt zwischen zwei Strecken S_1 und S_2 liegt, welche durch die vier Punkte S_{1a}, S_{1b}, S_{2a} und S_{2b} gegeben sind, liegt nach einer projektiven Transformation

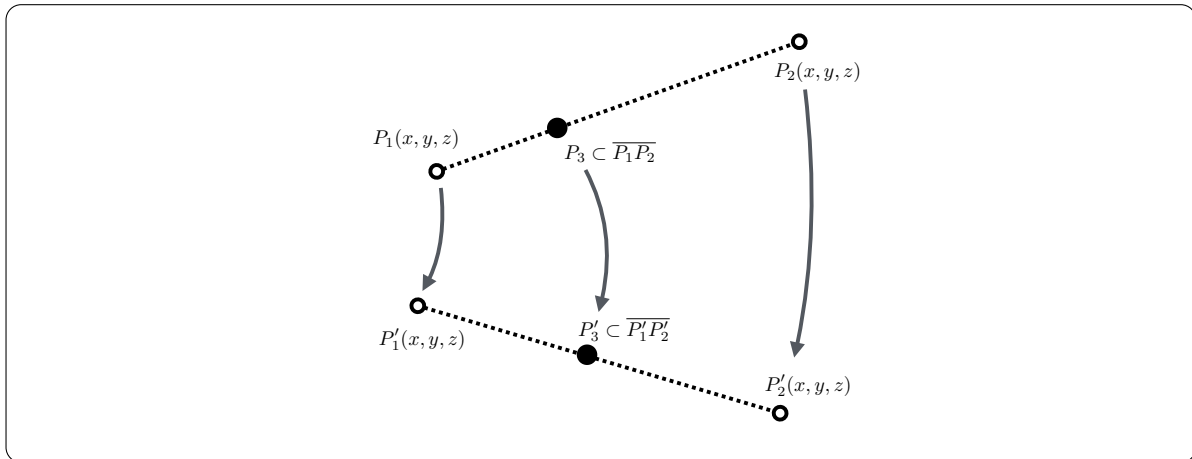
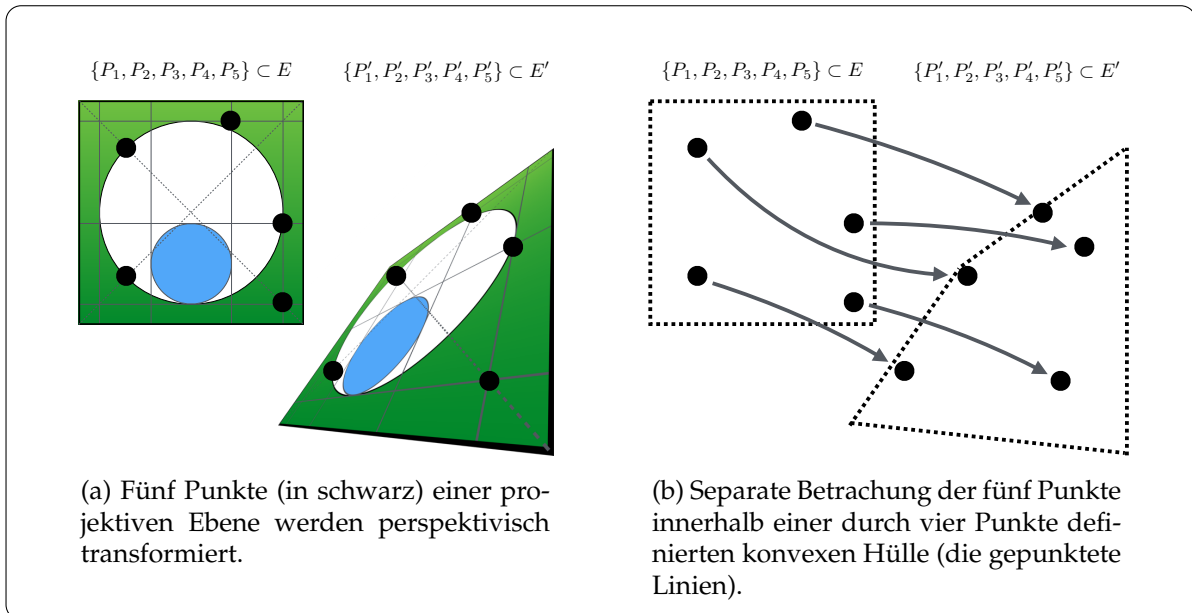


Abbildung 3.2: Geometrische Inzidenz: Auftreten des Punktes P_3 auf der Strecke $\overline{P_1P_2}$. Wenn ein Punkt innerhalb der Strecke zwischen zwei räumlichen Koordinaten liegt, so liegt dieser Punkt auch nach einer perspektivischen Transformation noch innerhalb derselben transformierten Strecke.



(a) Fünf Punkte (in schwarz) einer projektiven Ebene werden perspektivisch transformiert.
 (b) Separate Betrachtung der fünf Punkte innerhalb einer durch vier Punkte definierten konvexen Hülle (die gepunktete Linien).

Abbildung 3.3: Geometrische Inzidenz: Auftreten mehrerer Punkte P_1, P_2, P_3, P_4 und P_5 innerhalb der Ebene E . Wenn ein Punkt in Ebene E auftritt, so tritt dieser Punkt auch nach einer perspektivischen Transformation $(P'_1, P'_2, P'_3, P'_4, P'_5)$ noch auf der transformierten Ebene E' auf. Liegt ein Punkt innerhalb einer konvexen Hülle, so liegt der Punkt auch nach einer perspektivischen Projektion innerhalb dieser transformierten Hülle.

immer noch auf dem Schnittpunkt der beiden transformierten Strecken. Wie beispielhaft in Abbildung 3.3 zu sehen ist, bleiben die fünf schwarz markierten Punkte auch nach der projektiven Transformation auf dem jeweiligen Schnittpunkt der Linien, die sich bei den Punkten schneiden. Die Anordnung der Punkte ist hier so gewählt, dass maximal zwei Punkte kollinear sind, d.h. dass keine drei Punkte auf einer Geraden liegen. Auch nach der Transformation liegen keine drei Punkte auf einer Geraden im Bild. Auf diese Weise werden mit Inzidenzen Vorkommen von Punkten innerhalb von anderen Geometrien angegeben. Das ermöglicht es, die räumliche Anordnung von Punkten relativ zueinander anzugeben.

3.1.3 Projektiv planar invariante Merkmale

Das Doppelverhältnis von vier Punkten entlang einer projektiven Geraden

Das Doppelverhältnis von vier Punkten A, B, C und D entlang einer projektiven Geraden ergibt sich aus den Teilstrecken $[AC], [BC], [BD]$ und $[AD]$, wie in Gleichung 3.1 dargestellt.

$$\text{Doppelverhältnis}(A, B, C, D) = \frac{\frac{[AC]}{[BC]}}{\frac{[AD]}{[BD]}} \quad (3.1)$$

Das Doppelverhältnis ist projektiv invariant, was bedeutet, dass vier Punkte entlang einer projektiven Geraden sowohl im Ausgangsbild als auch im transformierten Bild dasselbe Doppelverhältnis besitzen. In Abbildung 3.4 ist das Doppelverhältnis schematisch dargestellt. Zu sehen sind vier Punkte auf der Diagonalen des Ausgangsbildes. Nach der Transformation des Bildes liegen die vier Punkte immer noch auf derselben Geraden, da sie in dieser enthalten sind (siehe Kapitel 3.1.2). Zu sehen ist auch, dass sich die Abstände zwischen den einzelnen Punkten im Bild stark verändern (Abbildung 3.4 b). Trotz dieser Veränderung der einzelnen Strecken zwischen den Punkten bleibt das Doppelverhältnis gleich (Gleichung 3.2).

$$\frac{\frac{[AC]}{[BC]}}{\frac{[AD]}{[BD]}} = \frac{\frac{[A'C']}{[B'C']}}{\frac{[A'D']}{[B'D']}} \quad (3.2)$$

Fünf Punkte auf einer projektiven Ebene

Neben dem Doppelverhältnis von vier Punkten auf einer projektiven Geraden gibt es noch eine Reihe weiterer projektiv planar invarianter Merkmale (siehe Kapitel 2.2.2). Aus diesen werden in dieser Arbeit die zwei funktional unabhängigen Invarianten I_1 und I_2 von fünf Punkten auf einer projektiven Ebene benutzt.

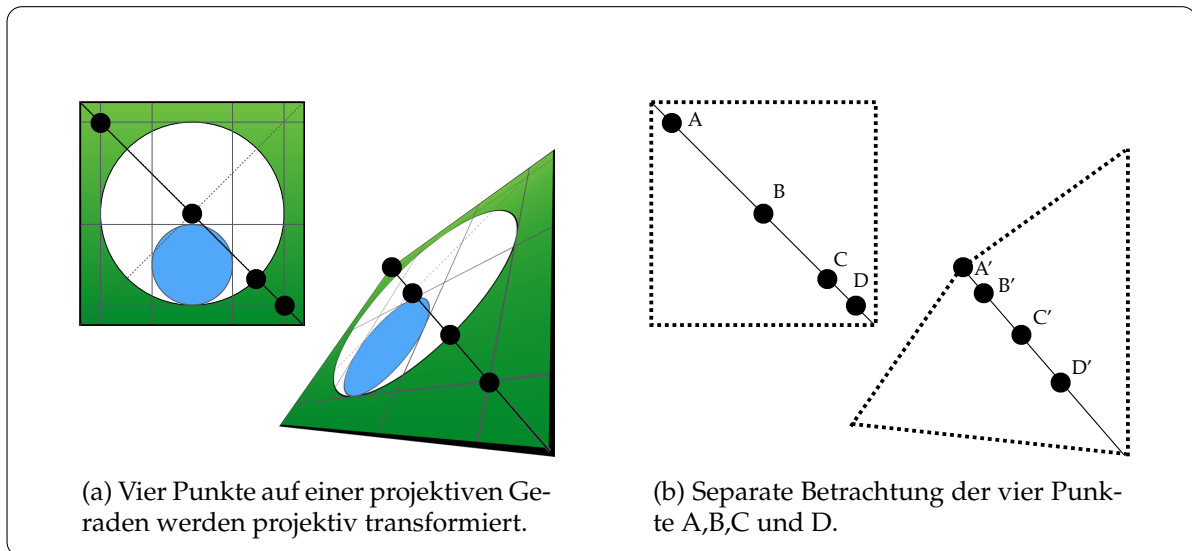


Abbildung 3.4: Bei einer projektiven Transformationen ändern sich die Abstände zwischen den einzelnen Punkten auf der Geraden.

Für fünf Punkte einer projektiven Ebene können zwei funktional unabhängige Invariante I_1 und I_2 konstruiert werden, solange keine drei der fünf Punkte kollinear sind [Mundy & Zisserman, 1992]. Die beiden Invarianten I_1 und I_2 werden aus den Bildkoordinaten der fünf Punkte errechnet. Die Formeln dafür können den Gleichungen 3.3, 3.4 und 3.5 entnommen werden. Diese Invariante der fünf Punkte auf einer projektiven Ebene wird im Folgenden als P^2 -Invariante bezeichnet.

$$m_{ijk} = (p_i, p_j, p_k) \text{ mit } p_i = (x_i, y_i, 1)^t \quad (3.3)$$

$$I_1 = \frac{|m_{431}| |m_{521}|}{|m_{421}| |m_{531}|} \quad (3.4)$$

$$I_2 = \frac{|m_{421}| |m_{532}|}{|m_{432}| |m_{521}|} \quad (3.5)$$

In Abbildung 3.5 sind die Konstruktionen der beiden Invarianten I_1 und I_2 anschaulich visualisiert.

3.1.4 Konnektivität

Eine weitere projektiv invariante Eigenschaft sind Konnektivitäten. Die Konnektivität beschreibt im Falle von Bildpunkten die Verbindungen und die Reihenfolge von Nachbarpunkten in der Umgebung eines Punktes. So bleiben von einem Punkt aus gesehen die Abfolge der Nachbarpunkte auf einer projektiven Ebene gleich. Ein

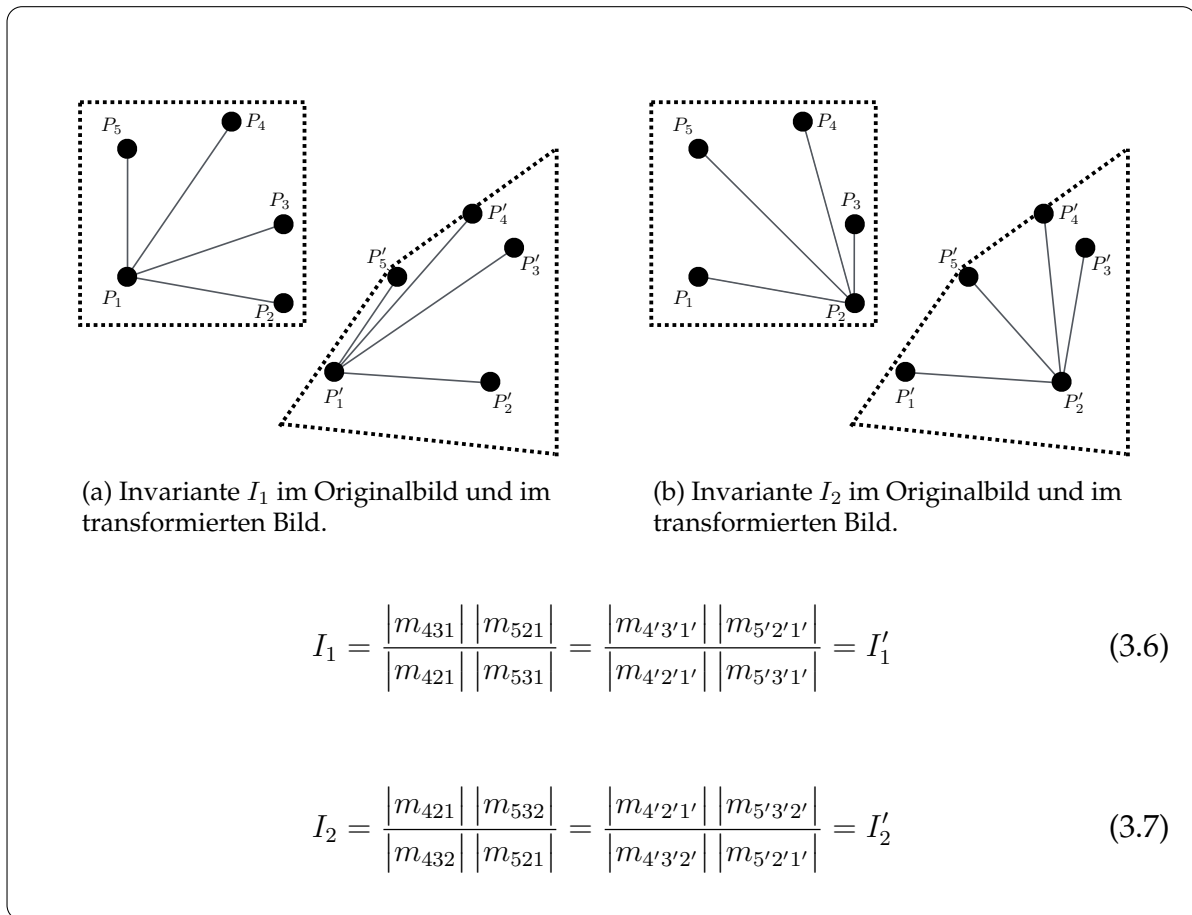


Abbildung 3.5: Die Invariante I_1 und I_2 von fünf identischen Punkten auf einer projektiven Ebene, sind im Originalbild und im transformierten Bild gleich.

Punkt P und seine Nachbarnpunkte N_1, N_2, N_3 und N_4 , welche auf derselben projektiven Ebene liegen, behalten ihre Reihenfolge gesehen von Punkt P aus bei. Wenn die Abfolge aus Sicht von P im Uhrzeigersinn zuerst N_1 , dann N_2 , danach N_3 und zuletzt N_4 ist, so ist diese Abfolge auch im projektiv transformierten Bild die gleiche, also N'_1, N'_2, N'_3, N'_4 (siehe Abbildung 3.6 a). Auch die Reihenfolge der Punkte auf der konvexen Hülle einer Punktmenge bleibt gleich. Ist die Reihenfolge der Punkte auf der konvexen Hülle im Ausgangsbild H_1, H_2, H_3, H_4, H_5 , dann ist auch die Reihenfolge der transformierten Punkte $H'_1, H'_2, H'_3, H'_4, H'_5$ auf der konvexen Hülle die gleiche (siehe Abbildung 3.6 b).

Wie oben ausgeführt wurde, beschreiben die projektiv invarianten Eigenschaften räumliche Zusammenhänge zwischen mehreren Punkten im Bild. Ein projektiv invariantes Merkmal, welches zur Bildbeschreibung benutzt wird, muss daher aus mehreren Punkten bestehen, deren Zusammenhänge durch die invarianten Kenngrößen definiert werden.

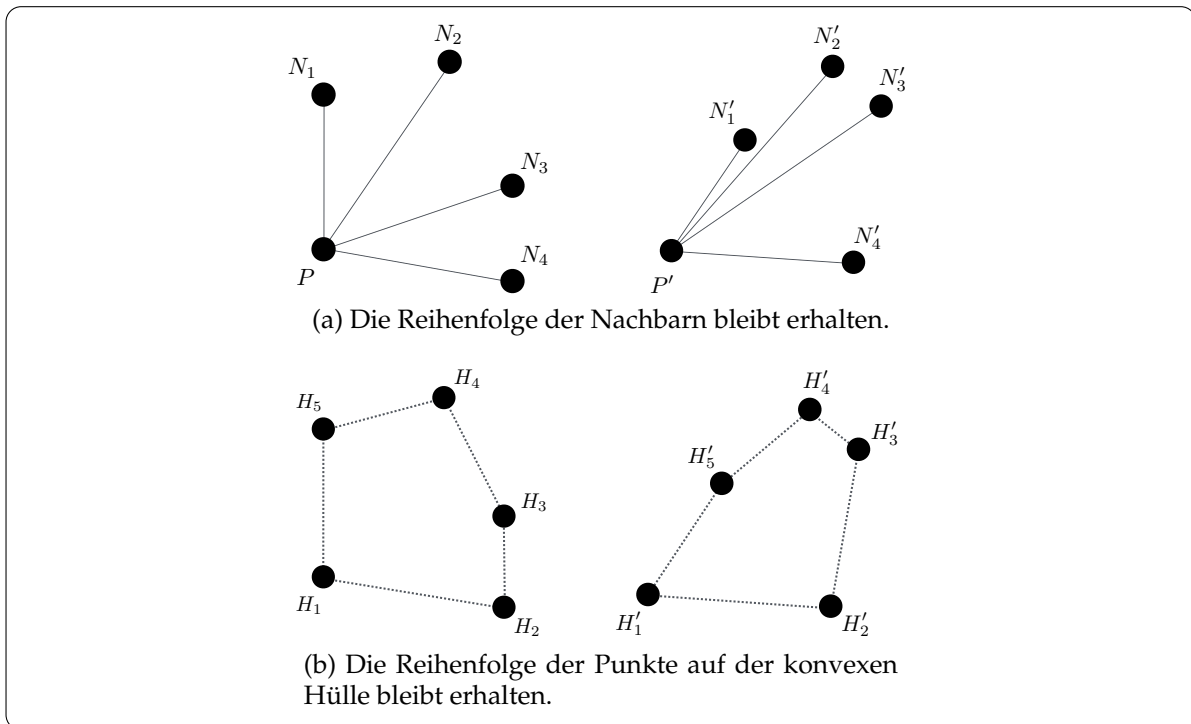


Abbildung 3.6: Die Konnektivität von Punkten im Bild bleibt bei projektiven Transformationen erhalten. Die Reihenfolge der Nachbarn von einem Punkt aus gesehen bleibt gleich und die Reihenfolge der Punkte auf der konvexen Hülle bleibt gleich.

3.2 Konzeption radiometrischer Invarianzen

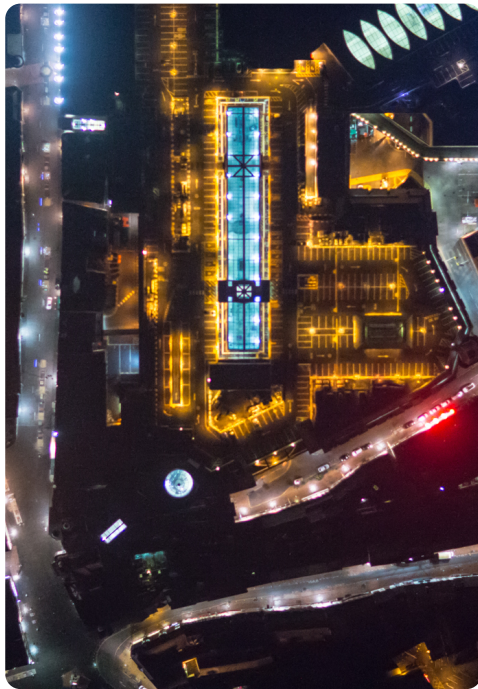
3.2.1 Ansprüche an die Merkmale

Die Herausforderungen, identische Merkmale in Bildern mit stark unterschiedlicher Radiometrie zu finden, wurden in Kapitel 2.3.1 angesprochen. Im Folgenden werden Bildmerkmale aufgeführt, die auch bei sich verändernder Radiometrie gleich bleiben. Dazu zählen Veränderungen der Beleuchtungsquellen, unterschiedliches Wetter, unterschiedliche Tages- und Nachtzeitpunkte einschließlich der daraus resultierenden unterschiedliche Schattenwürfe, unterschiedliche Zeitpunkte im Jahr sowie unterschiedliche Aufnahmesensoren.

Ein Beispiel mit sehr stark veränderter radiometrischer Information desselben Objektes ist in Abbildung 3.7 dargestellt. Die beiden Luftbildaufnahmen zeigen dasselbe städtische Gebiet in Leicester in England. Die linke Aufnahme ist eine Luftbildaufnahme bei Nacht, die rechte eine Luftbildaufnahme bei Tag. Der große Unterschied, der zu sehen ist, sind die unterschiedlichen Farben der Beleuchtungen, einmal durch

3 Grundlagen und Konzeption

Lampenlicht und einmal durch die Sonne. Bei der Nachtaufnahme erscheinen beispielsweise die Straßenlaternen in Weiß oder die Parkplatzbeleuchtung lässt die Parkplatzmarkierungen in Gelb erscheinen, wo hingegen bei der Luftbildaufnahme die Straßenlaternen nicht sichtbar sind und die Parkplatzmarkierungen weiß sind. Eine weiterer offensichtlicher Unterschied ist, dass in der Nachtaufnahme große Bereiche auf den Dächern der Gebäude unbeleuchtet sind und daher schwarz erscheinen. Genau umgekehrt dazu sind in der Luftbildaufnahme bei Tag die Dächer gut sichtbar und die Straßenzüge liegen im Schatten.



(a) Luftbildaufnahme eines Gebäudekomplexes mit einem Parkplatz auf dem Dach bei Nacht.



(b) Gebiet wie (a) bei Tag aufgenommen.

Abbildung 3.7: Starke Unterschiede in der radiometrischen Bildinformation des gleichen Objektes. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)

Es ist offensichtlich, dass ein Mensch trotz aller Unterschiede schnell identische Punkte findet, und das auch, wenn er keine weiteren Modellinformationen zur Verfügung hat und lediglich die beiden Bilder sieht. Die Suche nach identischen Bildpunkten kann daher nicht über die Farbinformation an sich erfolgen, da diese in beiden Bildern zu stark voneinander abweicht. Was allerdings gleich bleibt, sind die Positionen der deutlichen Farbänderungen, also die Kanten im Bild. Auf dem Dachparkplatz des großen Gebäudes sind die Linien der Parkplatzumrandungen im Nachtbild und im Tagbild gut sichtbar und manuell zuordenbar. In beiden Fällen liegt hier eine Ver-

änderung der Helligkeitsinformation vor. Auch die mandelförmigen Glaslichtdurchlässe des Gebäudes am rechten oberen Bildrand sind eindeutig zuordenbar, obwohl sich die Helligkeitsinformation im Nachtbild genau umgekehrt verhält wie im Tagbild. Bei Nacht erscheinen die aus Glas gebauten Lichtdurchlässe heller als die Umgebung, bei Tag hingegen dunkler. Die Position der Veränderungen der Intensitätsinformation allerdings bleibt gleich. Die Eigenschaft der gleichbleibenden Kanten wird im Folgenden genauer betrachtet.

3.2.2 Kanten als Inzidenzen

Kantenbilder beschreiben die Veränderungen der Intensitätsinformationen der einzelnen Kanäle im Bild. Wenn beispielsweise ein sehr niedriger Intensitätswert neben einem deutlich höheren Intensitätswert in einer Bildmatrix gespeichert sind, dann entsteht im Bild eine Kante. Kantendetektoren suchen auf unterschiedliche Weise nach diesen Sprüngen im Bild und markieren diese. Der bereits 1968 veröffentlichte Sobel-Operator [Sobel & Feldman, 1968] beispielsweise sucht mit einer kleinräumigen Maske mit vergleichsweise einfachen Rechenoperationen nach Intensitätssprüngen im Bild. Der robuste Canny Kantendetektor [Canny, 1986] benutzt einen deutlich aufwändigeren mehrstufigen Algorithmus. Bei letzterem wird nicht jeder Gradient im Bild angezeigt, sondern nur deutlich sichtbare und markante Kanten.

In Abbildung 3.8 sind die Ergebnisse beider Algorithmen für einen kleinen Ausschnitt aus den Luftbildern aus Abbildung 3.7 zu sehen. Auch wenn die radiometrische Information der ursprünglichen RGB-Bilder aufgrund der unterschiedlichen Tages- und Nachtzeitpunkte sehr unterschiedlich ist, fällt die Übereinstimmung von Teilen der Kantenbilder sofort ins Auge. Diese Eigenschaft, dass viele Kanten trotz sich verändernder Radiometrie erhalten bleiben, wird im Folgenden zur radiometrisch invarianten Beschreibung der Bildinformation benutzt.

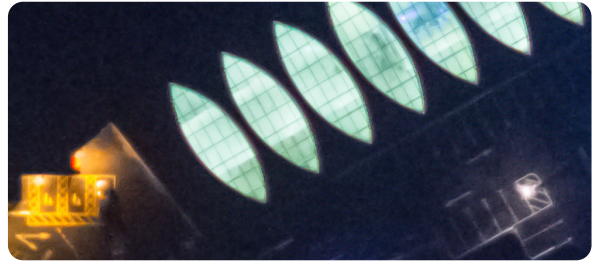
3.2.3 Eckpunkte als Inzidenzen

Neben den Kanten können Eckpunkte als radiometrisch invariante Merkmale benutzt werden. Dies funktioniert besonders gut, wenn diese Eckpunkte auch gleichzeitig physikalische Eckpunkte einer Geometrie darstellen. So bildet sich beispielsweise die Ecke einer Fensternische unabhängig von der Beleuchtung und der Blickrichtung in einem anderen Bild ebenfalls als Ecke der gleichen Fensternische ab. Weil die Kante eines Tiefsprungs eine Kante bleibt und die Ecke entlang dieser Kante auch eine Ecke bleibt, können geometrische Kanten und Ecken als projektiv und radiometrisch invariante Identifizierungsmerkmale benutzt werden. Viele unterschiedliche Algorithmen bieten die Möglichkeit, Eckpunkte in Bildern zu erkennen. In dieser Arbeit werden Harris Eckpunkte [Harris & Stephens, 1988] und SFOP-Junctions

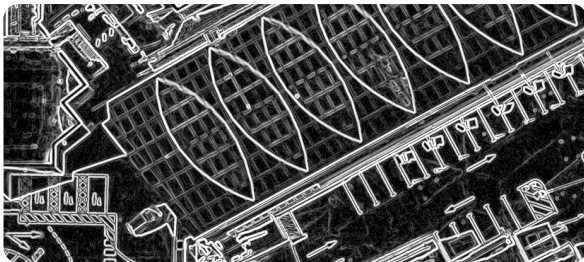
3 Grundlagen und Konzeption



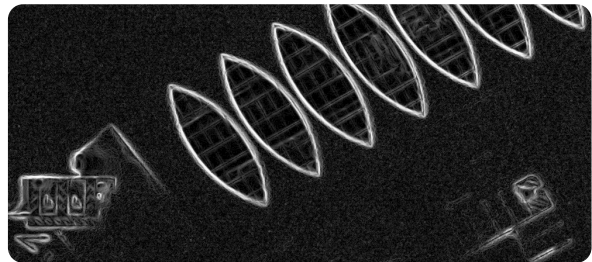
(a) Luftbildaufnahme bei Tag



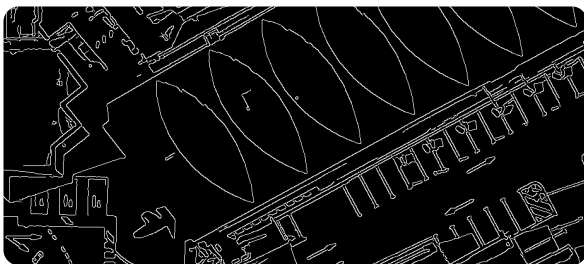
(b) Luftbildaufnahme bei Nacht



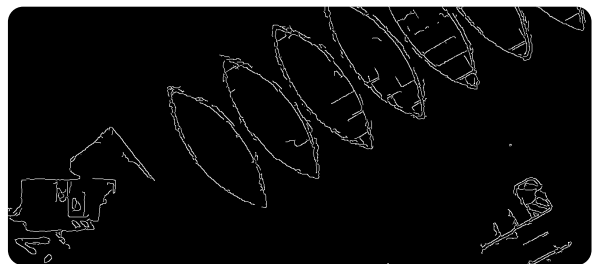
(c) Kantenbild durch Sobel-Operator von Grauwertbild von Bild a



(d) Kantenbild durch Sobel-Operator von Grauwertbild von Bild b



(e) Kantenbild durch Canny Kantendetektor von Bild a



(f) Kantenbild durch Canny Kantendetektor von Bild b

Abbildung 3.8: Gegenüberstellung der ursprünglichen RGB-Bilder und der Kantenbilder aus zwei unterschiedlichen Algorithmen. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)

[Förstner *et al.*, 2009] benutzt. Harris Eckpunkte werden anhand der lokalen Maxima in den Gradientenbildern in der unmittelbaren Umgebung eines Punktes gefunden. SFOP-Junctions verwenden neben den Gradienten in der Umgebung der Punkte auch Linienschnitte, die daraus errechnet wurden. Dies ermöglicht eine robuste skaleninvariante Beschreibung von Eckpunkten. In Abbildung 3.9 sind exemplarisch SFOP-Junction Eckpunkte an einem Tages- und einem Nachtluftbild dargestellt. Die unterschiedlichen Auswirkungen der beiden Algorithmen werden in Kapitel 6.2.1 besprochen.

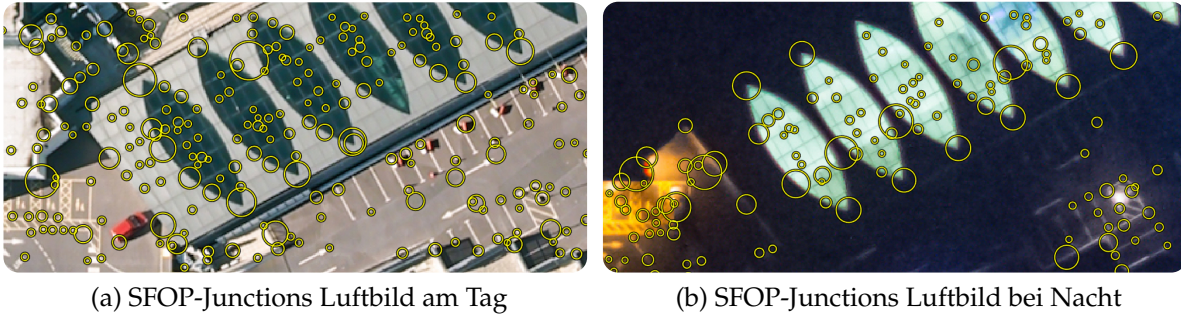


Abbildung 3.9: Beispiele für radiometrisch invariante Eckpunkt-Features. Nicht alle Eckpunkte stimmen im Tag- und im Nachtluftbild überein. Entscheidend ist jedoch, dass einige der Punkte übereinstimmen, die an markanten Stellen in beiden Bildern detektiert werden. Dies trifft zum Beispiel auf die Ecken an den Enden der großen mandelförmigen Glasfenster zu. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)

3.3 Anwendung bekannter projektiv invarianter Flächenmerkmale

[Meer *et al.*, 1994, 1998] beschreiben, wie mit Hilfe von P^2 -Invarianten (siehe Kapitel 3.1.3) manuell gemessene Bildpunkte einander zugeordnet werden können, wenn diese in einer Ebene liegen. Das neue Konzept, welches in der vorliegenden Arbeit entwickelt wurde, verwendet für die durchgeführten Plausibilitätstests (siehe Kapitel 5.2.2) mehrere Prinzipien dieser Methode. Daher wird diese Herangehensweise im Folgenden genauer beschrieben, sowie deren Vor- und Nachteile erörtert.

3.3.1 Flächenmerkmale mit Hilfe von P^2 -Invarianten

Die Methode, P^2 -Invarianten dazu zu benutzen, um manuell gemessene Bildpunkte anhand ihrer Koordinaten einander zuzordnen wird als *blinde Technik* bezeichnet, da sie außer den verwendeten Punktkoordinaten keine weiteren Informationen benutzt [Meer *et al.*, 1994]. Der Lösungsansatz ist ein Brute-Force Ansatz, bei dem alle Punktkombinationen zwischen den gemessenen Bildpunkten in beiden Bildern untersucht werden. Zur Verifizierung einer Punktekombination wird die P^2 -Invariante benutzt. Dazu werden innerhalb eines Bildes von jedem untersuchten Punkt aus alle Fünferkombinationen zu allen anderen Punkten im Bild benutzt. Die Anzahl an Kombinationen ergibt sich dabei aus dem Binomialkoeffizient (Formel 3.8), wobei n die Anzahl der Punkte im Bild ist und $k = 5$ ist.

$$\frac{n!}{k!(n-k)!} = \binom{n}{k} \quad (3.8)$$

Beispielsweise ergeben sich für 20 Bildpunkte 15 504 Kombinationsmöglichkeiten innerhalb eines Bildes. Alle Fünferkombinationen an einem Punkt im ersten Bild werden nun mit allen Fünferkombinationen eines Punktes im zweiten Bild verglichen, was bei 20 Bildpunkten pro Bild 240 374 016 Kombinationsmöglichkeiten zwischen beiden Bildern entspricht. Dabei werden jeweils die P^2 -Invarianten I_1 und I_2 (siehe Kapitel 3.1.3) fünf mal berechnet und dann verglichen, also jeweils I_1 und I_2 an jedem der fünf Punkte. Auch wenn [Mundy & Zisserman, 1992] gezeigt haben, dass bei fünf Punkten nur ein I_1 und ein I_2 funktional unabhängig sind, so sind deren Werte komplett unterschiedlich, je nachdem von welchem Punkt der benötigten fünf Punkte aus die Berechnung gestartet wird. Da aufgrund der *blinden Technik* nicht bekannt ist, von welchem Punkt aus gezählt wird, werden alle fünf Punkte als Startpunkte benutzt, was den Rechenaufwand weiter erhöht.

Stimmen I_1 und I_2 überein, so wird die Anzahl der Punkte auf der konvexen Hülle der fünf Punkte überprüft, ob sie in beiden Bildern die gleiche ist. Danach wird zusätzlich die Reihenfolge der Punkte auf der konvexen Hülle verglichen. Stimmen auch diese überein, wird das Kandidatenpaar in einer Liste mit dem Vorkommen=1 eingetragen. Sollte das Vorkommen aus einer anderen Kombination bereits eingetragen sein, so wird der Zähler um eins erhöht. Auf diese Weise wird für jede Kombination an Bildpunkten aus beiden Bildern überprüft, wie oft Übereinstimmungen in beiden Bildern gefunden wurden. Abschließend werden die Punkte, welche die meisten Übereinstimmungen aufweisen, als Punktzuordnung ausgewiesen. Prinzipiell könnte bereits nach der ersten gefundenen Übereinstimmung zwischen zwei Kandidaten dieser Kandidat als zugeordnet angesehen werden. Dies würde funktionieren, wenn die Genauigkeit von I_1 und I_2 maximal hoch ist, also sehr viele Dezimalstellen berücksichtigt werden. Aufgrund der Punktungenauigkeiten der gemessenen Koordinaten der Bildpunkte und der Verzeichnung der Bilder bei Realdaten können diese aber nur mit einer geringen Genauigkeit verwendet werden (siehe [Mundy & Zisserman, 1992] Seite 95). Durch diese geringere Genauigkeit fallen Mehrdeutigkeiten an, die allerdings durch die Gesamtzahl an gefundenen Übereinstimmungen aufgedeckt werden.

Mit dieser Methode ist eine 1:1 Punktzuordnung ebener Punkte möglich, wenn in beiden Bildern die gleiche Anzahl an Punkten vorliegt, oder wenn die Punkte des einen Bildes eine exakte Teilmenge der Punkte des anderen Bildes darstellen. Die Ergebnisse sind korrekt und es können selbst starke Blickwinkelveränderungen berücksichtigt werden. Der große Nachteil dieser Methode ist jedoch der enorm hohe Rechenaufwand bereits für kleine Punktmengen (siehe Beispiel in den Ergebnissen in Kapitel 6.5.2). Ein weiterer Nachteil ist, dass lediglich der am häufigsten zugeordnete Punkt als Korrespondenz ausgegeben wird. Sollte ein Punkt keine Entsprechung haben, wird er trotzdem eine Zuordnung eines Punktes des anderen Bildes erhalten.

3.3.2 Spezifische und eingeschränkte Verbesserungsansätze

Aus den Nachteilen, insbesondere dem extrem hohen Rechenaufwand ergeben sich folgende Ideen für Verbesserungsansätze. Um den Rechenaufwand zu verringern ist es notwendig, dass nicht alle Punkte mit allen anderen in den Fünferkombinationen kombiniert werden. Zudem ist es zielführend, Merkmale zu erstellen, die eine Identifizierung der benutzen 5 Punkte innerhalb einer Fünferkombination ermöglicht.

Eine Einschränkung auf beispielsweise räumlich naheliegende Nachbarn klingt im ersten Moment sehr plausibel. Sollen beliebige Veränderungen des Blickwinkels zugelassen werden, ist dies allerdings aus mehreren Gründen nicht allgemein gültig. Im Folgenden ist ein bebildertes Gegenbeispiel aufgeführt. Bereits aus Tabelle 2.1 geht hervor, dass Entfernungen und Schwerpunkte nicht projektiv invariant sind. Somit können auch keine Algorithmen, die diese Eigenschaften verwenden, benutzt werden. Dass diese Eigenschaften nicht projektiv invariant sind, wird gut sichtbar in Abbildung 3.10 dargestellt.

Alle projektiv invarianten Eigenschaften geben Beziehungen zwischen mehreren Einzelpunkten an (vergleiche Tabelle 2.1). Somit können keine rein auf Einzelpunkte bezogenen Identifizierungsmerkmale erstellt werden, die projektiv invariant sind. Um dennoch mit einer minimalen Anzahl an Punkten unterscheidbare Merkmale zu erstellen, werden die in Kapitel 4 beschriebenen virtuellen Linienmerkmale benutzt, welche die wichtigste Neuheit dieser Arbeit sind.

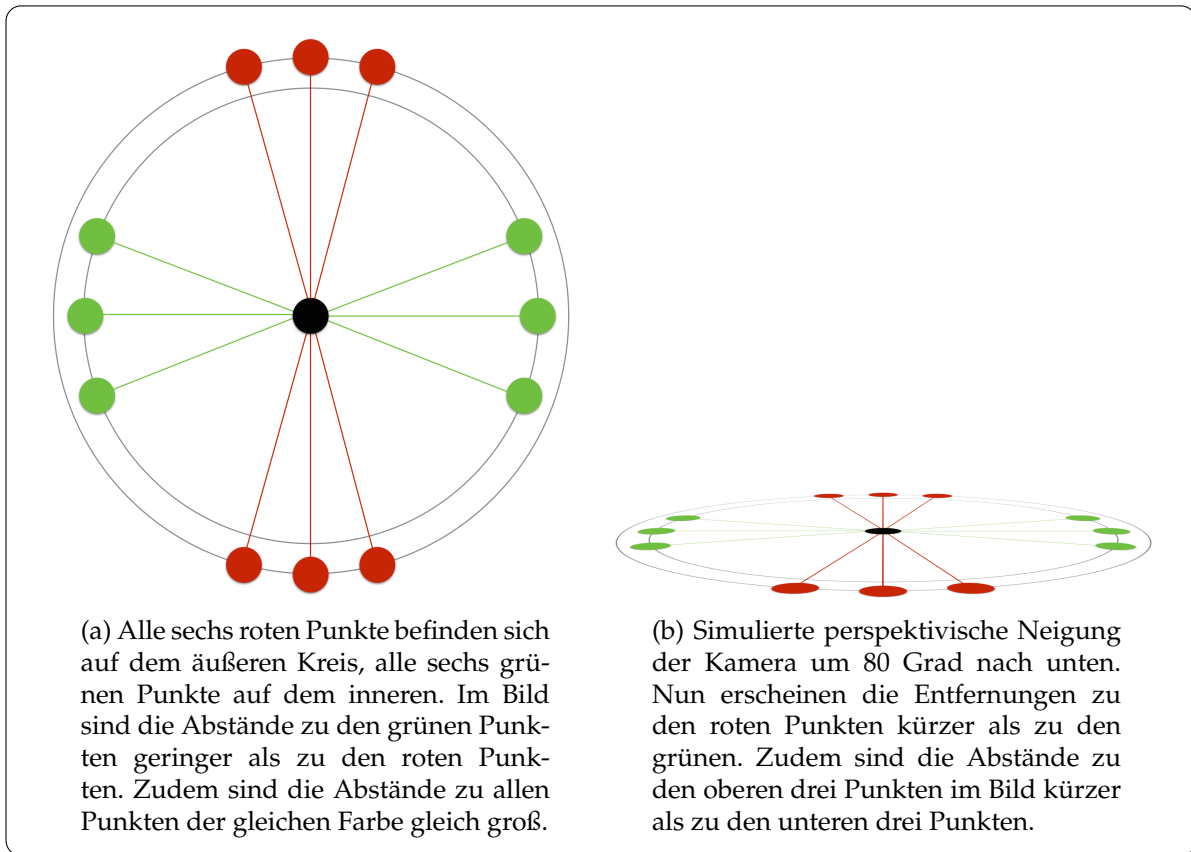


Abbildung 3.10: Eine Verdeutlichung, warum Vermaschungen auf Grundlage von nächsten Nachbarschaften und Schwerpunkten nicht perspektivisch invariant sind. Ein Algorithmus, der die nächsten sechs Punkte abfragt, gibt im linken Bild die grünen Punkte zurück, im rechten, um 80 Grad geneigten Bild, die roten Punkte.

4 Neues Konzept der Zuordnung virtueller Linienmerkmale

4.1 Grundidee

Die in dieser Arbeit verwendeten virtuellen Linienmerkmale beschreiben Linien, die nicht als erkennbare Linien im Bild sichtbar sind. Im Gegensatz zu sichtbaren Linien, die aus Bildern extrahiert werden können, wie beispielsweise bei der Extraktion von Straßennetzen [Hinz, 2003] oder Passgeraden im Objektraum zur relativen Bildorientierung [Strunz, 1993], sind die in diesem Kapitel vorgestellten Linien nicht direkt im Bild interpretierbar. Es handelt sich um gerade Linien, die künstlich zwischen markanten Eckpunkten erstellt worden sind. Es werden hier ausschließlich Methoden angewandt, die bei einer starken projektiven Blickwinkelveränderung Bestand haben und die in Kapitel 3.1 vorgestellt wurden.

Eine mit dem Themengebiet verwandte ausgezeichnete Arbeit, die auch auf virtuellen Linienmerkmalen beruht, sind die sogenannten *Descriptor-Nets* ([von Hundelshausen & Sukthankar, 2012]). Deren Methode benutzt Pixelintensitäten entlang der Linien zur Merkmalsbeschreibung. Leider ist dieser Algorithmus nur invariant gegenüber affinen Transformationen, was unter anderem daran liegt, dass intern eine Delaunay Triangulation benutzt wird, um benachbarte Punkte zu finden. Diese ist wie angesprochen nicht projektiv invariant, da bei der Triangulation entfernungsabhängige Nachbarschaften benutzt werden (vergleiche Abbildung 3.10). Auch wenn die *Descriptor-Nets* nicht projektiv invariant sind, so zeigt der Ansatz dennoch sehr gut, dass auch virtuelle Linienmerkmale hervorragend für Bildbeschreibungen verwendbar sind.

Der Ansatz von [Luo *et al.*, 2013] findet wie in Kapitel 2.2.2 beschrieben mit Hilfe eines *Cross-Ratio-Spektrum* abgeschlossene Symbole, die projektiv transformiert wurden. Intern werden auch hier Linienmerkmale bei der Erstellung des Features benutzt. Das Auffinden von identischen Bildbereichen erfolgt jedoch im Gegensatz zu der hier vorgestellten Methode Featurebasiert über das *Cross-Ratio-Spektrums*. Trotz der Beschränkung des Algorithmus auf ideale 1:1 Zuordnungen, und dass infolge dessen keine Teilbilder oder Verdeckungen identifiziert werden können, kann dieser Ansatz als der bisher vielversprechendste gesehen werden, da er komplett projektiv invariant ist, was in [Fan *et al.*, 2014] gezeigt wurde.

4 Neues Konzept: Linienmerkmale und deren Zuordnung

Im Folgenden wird das neue Konzept der projektiv invarianten virtuellen Linienmerkmale und deren Zuordnung vorgestellt. Die Grundidee ist es, virtuelle Linien im Bild zwischen markanten Punkten (siehe Abbildung 4.1) zu erstellen und diese Linien mit Linien in anderen Bildern zu vergleichen. Entlang dieser Linien werden



(a) Eckpunkte im ersten Bild (exemplarisch)



(b) Eckpunkte im zweiten Bild (exemplarisch)

Abbildung 4.1: Als Markante Punkte im Bild werden Eckpunkte benutzt. Hier sind exemplarisch einige der Eckpunkte mit roten Punkten im Bild markiert.

zunächst die Schnittpunkte mit darunter liegenden Kanten im Bild detektiert. Die Positionen der Schnittpunkte entlang der Linien werden mit Hilfe von Doppelverhältnissen projektiv invariant beschrieben. Ein Linienfeature besitzt ein Doppelverhältnis oder mehrere Doppelverhältnisse als Merkmale. Diese Doppelverhältnisse beschreiben die Positionen der Schnittpunkte der Linie mit dem Kantenbild. Die Idee dahinter ist, dass sich nach einer projektiven Transformation zwar die Positionen der Gradientenschnitte innerhalb einer Linie verändern, deren Doppelverhältnisse aber gleich bleiben, wenn die Linie und die Gradienten tatsächlich die gleichen sind und in einer Ebene liegen (siehe Abbildung 4.2). Eine weitere neue Idee ist es, dass nicht alle Doppelverhältnisse der Gradienten gleich sein müssen, aber mindestens drei. Dies indiziert, dass mindestens drei der Gradientenschnitte übereinstimmen ($\binom{3}{2} = 3$). So können Verdeckungen oder fehlende Bildinformationen trotzdem korrekt wiedererkannt werden, solange wenigstens ein Teil der Information vorhanden ist. Dies ist vor allem in Kombination mit den projektiv invarianten Flächenmerkmalen (siehe Kapitel 5) wichtig, da mit diesem Konzept auch die Wiedererkennung von Teilbereichen realisiert wird, was bei [Luo *et al.*, 2013] und [Fan *et al.*, 2014] nicht möglich ist. Eine weitere Neuheit ist die punktweise Zuordnung der Linienfeatures, die in Kapitel 4.3 beschrieben wird.

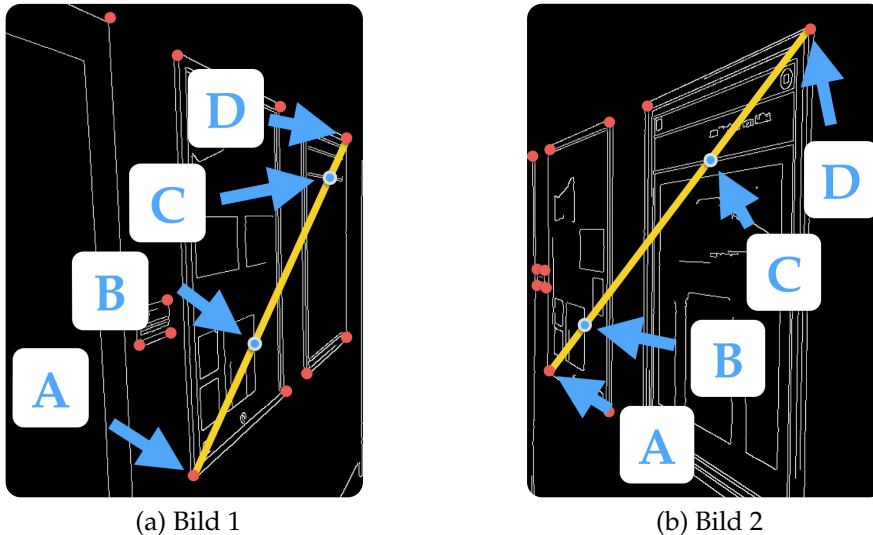


Abbildung 4.2: Zwischen den Eckpunkten werden die virtuellen Linienmerkmale erstellt. Die Schnittpunkte der Linie mit dem Kantenbild B und C bilden zusammen mit den Anfangs- und Endpunkten der Linie A und D vier Punkte auf einer geraden Linie. Von diesen vier Punkten kann nun das Doppelverhältnis ausgerechnet werden. Wenn die Punkte in beiden Bildern dieselben sind, so ist auch dieses Doppelverhältnis gleich.

4.2 Merkmalsbeschreibung bei Linienmerkmalen

4.2.1 Eckpunkte und Kanten als Eingangsdaten

Die Vorgehensweise zur Erstellung von Linienfeatures (siehe auch Abbildung 4.4) ist folgende. Zunächst werden mit Hilfe eines Eckpunktdetektors Eckpunktfeatures aus dem Bild extrahiert. Hierfür werden beispielsweise Harris- oder Förstner-Features verwendet [Harris & Stephens, 1988; Förstner & Gülch, 1987; Förstner *et al.*, 2009], siehe auch Kapitel 3. Danach werden die markanten Kanten im Bild mit Hilfe des Canny-Algorithmus [Canny, 1986] detektiert und als Bildmatrix gespeichert. Die Bildkoordinaten der Eckpunkte und die Kantenbilder sind die Eingangsdaten zur Erstellung der projektiv invarianten Linienfeatures.

4.2.2 Erstellung der Linienmerkmale

Innerhalb eines Bildes werden zwischen allen Eckpunkten virtuelle Linien erstellt. Entlang jeder Linie werden für die darunter liegenden Pixel des Kantenbildes die jeweiligen Werte abgefragt. Im Fall von Canny-Kantenbildern kann der Wert 0 oder 1 sein, also keine Kante oder eine Kante. Die Distanzen zwischen den Kantenpixeln

4 Neues Konzept: Linienmerkmale und deren Zuordnung

und dem Anfangspunkt der Linie werden je Linie gespeichert. Somit wird für jede Linie eine Liste erstellt, in der die Entfernungen der darunter liegenden Kantenpixel zum Anfangspunkt der Linie stehen. Um ein projektiv invariantes Doppelverhältnis ausrechnen zu können (siehe Kapitel 3.1.3), werden vier Punkte entlang einer Geraden benötigt, um daraus die dazwischen liegenden Strecken zu errechnen. Anfangs- und Endpunkt einer Linie sind jeweils zwei dieser Punkte. Die anderen zwei Punkte werden durch die Kantenpixel entlang der Linien vorgegeben. Deshalb wird anschließend für jede Linie getestet, ob mindestens zwei weitere Schnittpunkte mit dem Kantenbild enthalten sind. Bei genau zwei weiteren Schnittpunkten kann ein Doppelverhältnis errechnet werden. Wenn mehr als zwei Kantenpixel in einer Linie enthalten sind, können alle Kombinationen für $\binom{n}{2}$ Doppelverhältnisse entlang der Linie benutzt werden, wobei n die Anzahl der Kantenschnittpunkte mit der virtuellen Linie ist (siehe Abbildung 4.3).

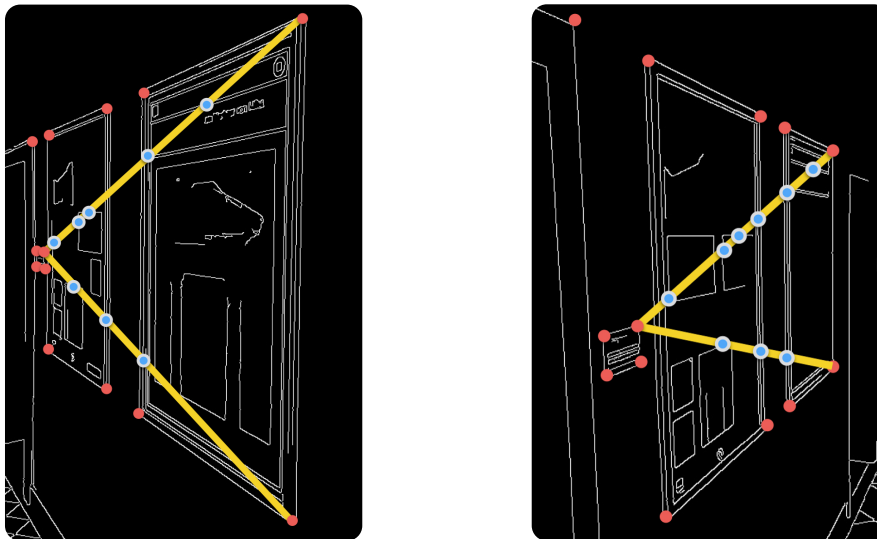


Abbildung 4.3: Beispiel von mehreren Doppelverhältnissen entlang der virtuellen Linien. Die Doppelverhältnisse entstehen durch den Linienanfangs- und Endpunkt, sowie durch die dazwischen liegenden Schnittpunkte mit dem Kantenbild.

Die Liste der Doppelverhältnisse einer Linie zwischen zwei Eckpunktfeatures enthält die projektiv invarianten Merkmale der virtuellen Linienfeatures. Die Richtung der Linie, also welches Ende der Anfangs- und welches der Endpunkt ist, ist bei dieser Repräsentation irrelevant. Das Doppelverhältnis der vier Teilstrecken ist unabhängig von der Richtung der Linie.

Da es vorkommen kann, dass entlang einer Linie viele Kanten sehr dicht aufeinander folgen, ergeben sich hieraus viele numerisch ähnliche Doppelverhältnisse. Um für das anschließende Featurematching möglichst eindeutige Features zu benutzen, werden die ähnlichen Doppelverhältnisse anhand ihres Abstandes im Vektorraum auf

aussagekräftige Doppelverhältnisse reduziert. Ein Mindestabstand von 0.1 hat sich hier als geeignet herausgestellt (siehe Kapitel 6.2.7). Höhere Mindestabstände führen zu kürzeren Rechenzeiten aber auch weniger Zuordnungen, kürzere hingegen zu längeren Rechenzeiten und mehr Zuordnungen (siehe Abbildungen 6.18 sowie 6.19). Der gesamte Ablauf der Erstellung von projektiv und radiometrisch invarianten Linienmerkmalen ist in Abbildung 4.4 in einem Diagramm wiedergegeben.

4.3 Zuordnung von Linienmerkmalen

4.3.1 Vorgehensweise

Bei der Zuordnung der Linienmerkmale sind grundsätzlich zwei unterschiedliche Vorgehensweisen denkbar. Zum einen können die Linienmerkmale aus zwei unterschiedlichen Bildern direkt miteinander verglichen werden und die höchste Übereinstimmung gesucht werden. Zum anderen können alle benutzten Eckpunkte verglichen werden, ob an ihnen ähnliche Linienfeatures anliegen. Die rein auf Doppelverhältnisse bezogene Zuordnung von Linienfeatures hat den Nachteil, dass die Linienerichtung, also welcher Punkt der Anfangs- und welcher Punkt der Endpunkt ist, in dieser allgemeinen Form nicht direkt ersichtlich ist, da die Doppelverhältnisse an sich keine Rückschlüsse darauf liefern. Auch die Tatsache, dass viele gleiche Doppelverhältnisse häufig vorkommen, verhindert eine rein merkmalsbezogene Zuordnung. Stattdessen werden nacheinander alle Punktkombinationen zwischen beiden Bildern untersucht, um Einzelpunkte mit vielen identischen Linienfeatures zu finden.

4.3.2 Selektion beteiligter Linienfeatures

Die Vorgehensweise ist daher folgende. Zunächst werden alle Kombinationen von Punkten zwischen dem ersten und dem zweiten Bild aufgestellt und durch sie wird iteriert. Um vergleichbare virtuelle Linienfeatures zu erhalten, werden je Eckpunkt die zugehörigen Linienfeatures anhand der beteiligten Anfangs- oder Endpunktkoordinaten ausgewählt. Zu dem Eckpunkt mit den Koordinaten x_n und y_n werden alle entsprechenden Linienfeatures geladen, an denen der Anfangs- oder Endpunkt x_n und y_n sind. Dies wird für die untersuchten Punkte in beiden Bildern durchgeführt, so dass eine Liste mit allen beteiligten Linien am Eckpunkt in Bild 1 und eine Liste mit allen beteiligten Linien am gerade untersuchten Eckpunkt in Bild 2 vorliegen.

Als nächstes wird ein Test durchgeführt, ob überhaupt Linienfeatures an beiden Punkten vorhanden sind, da es theoretisch möglich ist, dass ein Punkt mit Linienfeatures mit anderen Punkten verbunden ist, die keine Schnittpunkte mit dem Kantenbild aufweisen. Dies kann beispielsweise bei Einzelpunkten innerhalb einer homogenen Fläche vorkommen.

4 Neues Konzept: Linienmerkmale und deren Zuordnung

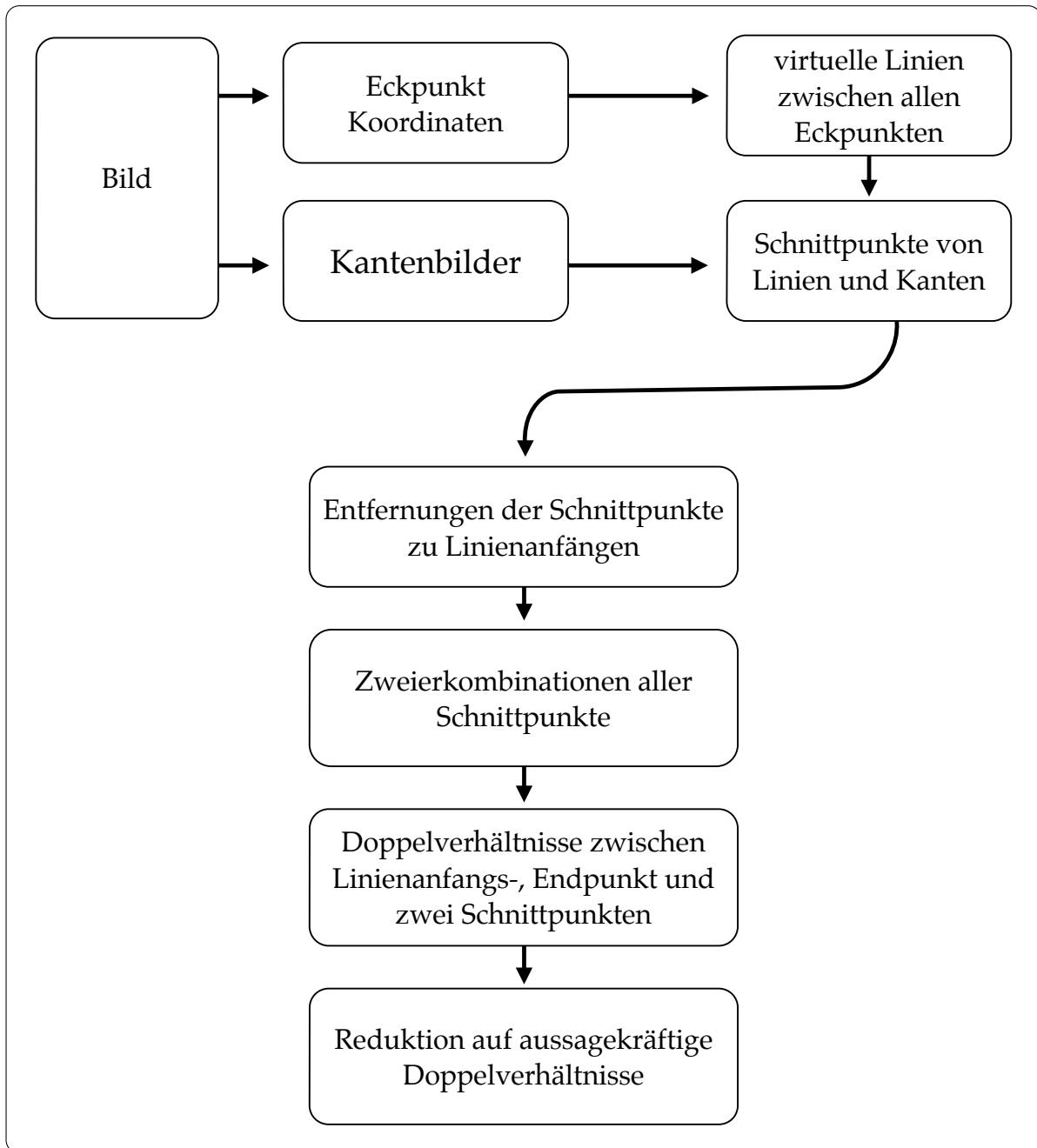


Abbildung 4.4: Der Aufbau der neu konzipierten projektiv und radiometrisch invarianten Linienmerkmale.

4.3.3 Vergleich von Linienfeatures anhand von Doppelverhältnissen

Wenn an beiden Punkten in den zwei Bildern Linienfeatures vorhanden sind, werden diese im nächsten Schritt miteinander verglichen. Dies erfolgt über einen einfachen Vergleich der Doppelverhältnisse (siehe Kapitel 3.1.3). Stimmen mindestens zwei Doppelverhältnisse in zwei Linienfeatures aus den zwei unterschiedlichen Bildern überein, so bedeutet dies, dass die Kanten zwischen zwei Eckpunkten möglicherweise an der gleichen Stelle liegen wie im anderen Bild. Je mehr Doppelverhältnisse übereinstimmen, desto größer ist das Indiz, dass hier eine gleiche Linie in beiden Bildern zwischen identischen Punkten vorliegt. Nach dem Vergleich aller an einem Punkt anliegenden Linienfeatures werden jene mit identischen Merkmalen zurückgegeben. Nun folgt ein weiterer Test, ob in beiden Bildern am untersuchten Punkt identische Features gefunden wurden.

4.3.4 Nachbarschaftsreihenfolge testen

Falls dies der Fall ist, werden im nächsten Schritt die Nachbarschaftsreihenfolgen der Linienfeatures in beiden Bildern verglichen (vergleiche Kapitel 3.1.4). Anhand der Merkmale (den Doppelverhältnissen) können die Linien in beiden Bildern unterschieden werden. Da sich oftmals nicht alle Merkmale unterscheiden, wird zunächst getestet, ob eindeutige Merkmale vorhanden sind. Nur wenn eindeutige Merkmale vorliegen, kann eine Überprüfung der Reihenfolge der Doppelverhältnisse durchgeführt werden. Da sich die Reihenfolge der Nachbarn bei gleichen Features auf einer Ebene nicht verändert, genügt ein einzelnes eindeutiges Linienmerkmal, von dem aus die Reihenfolge der Nachbarschaftspunkte getestet wird. Als Vergleichsgröße für die Reihenfolge der Nachbarn dienen die Richtungswinkel zwischen dem untersuchten Punkt und den Linienendpunkten. Die Richtungswinkel werden wie in Abbildung 4.5 beschrieben berechnet.

$$\text{Richtungswinkel} = \text{atan2}(y_{\text{Endpunkt}} - y_{\text{Kandidat}}, x_{\text{Endpunkt}} - x_{\text{Kandidat}}) \quad (4.1)$$

Abbildung 4.5: Berechnung des Richtungswinkels zwischen untersuchtem Punkt und Endpunkt des Linienfeatures.

Der Richtungswinkel des eindeutigen Linienmerkmals wird dabei als 0 definiert und alle anderen Richtungswinkel werden relativ dazu gemessen. Sollten die Endpunkte der Linienfeatures in beiden Bildern die gleichen sein, so müssten alle Richtungswinkel in beiden Bildern fortlaufend größer werden. Sollte dies nicht der Fall sein, wird getestet, bei wie vielen Nachbarn jeweils die Reihenfolge der Richtungswinkel stimmt, wenn unterschiedliche Startpunkte für den Nachbarschaftstest gewählt

werden. Die maximale Anzahl an Linienfeatures, die zu Nachbarn führen, bei denen die Reihenfolge als richtig erkannt wurde beim vorangegangenen Test, werden in Form einer Liste zurückgegeben. Inhalt dieser Liste sind die maximale Anzahl an Linienfeatures am untersuchten Punkt, deren Doppelverhältnisse in beiden Bildern übereinstimmen und deren Reihenfolge der Nachbarschaften übereinstimmen.

4.3.5 Ausgabe möglicher Punktkorrespondenzen

Auf diese Weise wird für jedes Kandidatenpaar der Punktkorrespondenzen eine Liste mit möglichen Nachbarn anhand der Linienfeatures erstellt. Die Anzahl an prognostizierten korrekten Nachbarpunkten kann bereits als Indikator für eine Punktzuordnung benutzt werden. Der Ablauf des gesamten Linienmatchings ist in Abbildung 4.6 zusammengefasst. Dieser rein auf virtuellen Linien basierende Ansatz wird im folgenden Kapitel 5 um weitere Plausibilitätstests aus Kapitel 3.3 ergänzt, um zuverlässigere Ergebnisse zu bekommen.

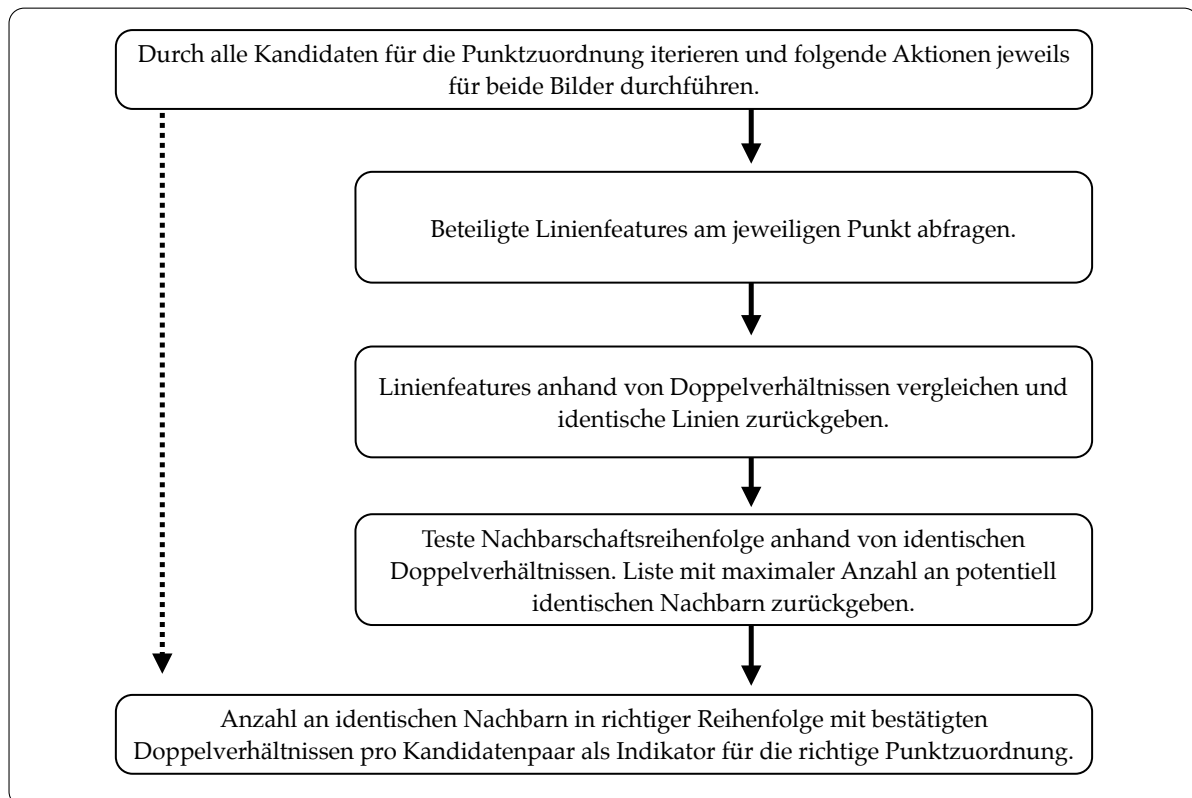


Abbildung 4.6: Konzept zum Auffinden von Punktkorrespondenzen mit den neuen virtuellen Linienfeatures.

5 Neues Konzept der kombinierten Linien- und Flächenmerkmale

Die in Kapitel 4 vorgestellten virtuellen Linienmerkmale können bereits dazu genutzt werden, mögliche Punktkorrespondenzen in zwei Bildern zu finden. Für den Fall, dass mehrere Punkte nur in einem Bild sichtbar sind, also beispielsweise bei Teilbildern, Bildausschnitten oder Verdeckungen und zur weiteren Verfeinerung des Algorithmus, werden weitere Tests mit den gefundenen Punktepaaren durchgeführt. Dabei werden auch hier ausschließlich Methoden angewandt, die bei einer projektiven Blickwinkelveränderung Bestand haben und die in Kapitel 3.1 vorgestellt wurden.

Eine Übersicht des gesamten Ablaufs von der Bildinformation der beiden Bilder über die virtuellen Linienmerkmale und die Plausibilitätstests bis hin zu den Punktkorrespondenzen ist am Ende dieses Kapitels in Abbildung 5.4 dargestellt.

5.1 Virtuelle Linienmerkmale für die initialen Hypothesen

Zunächst werden die gleichen Schritte wie in Kapitel 4 durchgeführt. Aus dem ersten und dem zweiten Bild werden jeweils die markanten Eckpunkte extrahiert und deren Koordinaten für die weitere Verarbeitung gespeichert. Ebenfalls aus beiden Bildern werden die Kantenbilder mit den markanten Kanten erstellt. Aus den Eckpunktkoordinaten und den Kantenbildern werden die beschriebenen neuen virtuellen Linienfeatures erstellt (vergleiche Abbildung 4.4). Für alle denkbaren Punktkorrespondenzen, welche sich aus den Eckpunktkoordinaten der beiden Bilder ergeben, wird eine Liste erstellt. Jeder Eintrag in dieser Liste steht für eine mögliche Punktkorrespondenz, die im Folgenden überprüft wird. Wie in Abbildung 4.6 beschrieben, wird für jedes Kandidatenpaar eine Liste mit der maximalen Anzahl an potentiell richtigen Nachbarn erstellt. Diese Liste ist der Ausgangspunkt für weitere Plausibilitätstests, die nachfolgend beschrieben werden.

5.2 Kombinierte Linien- und Flächenmerkmale

5.2.1 Test aller Kandidatenpaare

Die folgenden Tests können nur durchgeführt werden, wenn mindestens fünf Punkte involviert sind, da die P^2 -Invarianten I_1 und I_2 nur mit mindestens 5 Punkten ausgerechnet werden können (siehe Kapitel 3.1.3). Daher wird als erstes überprüft, ob in den vorausgegangenen Tests mindestens vier Nachbarpunkte gefunden wurden. Wenn dies der Fall ist, werden aus allen Nachbarpunkten Viererkombinationen gebildet, durch welche dann für die weiteren Plausibilitätstests iteriert wird, um diese Kombination entweder zu bestätigen oder zu verwerfen. Betrachtet werden immer vier Nachbarpunkte und der aktuell untersuchte Punkt aus der zu überprüfenden Punktkorrespondenz, wodurch jeweils fünf Punkte untersucht werden (siehe Abbildung 5.1).

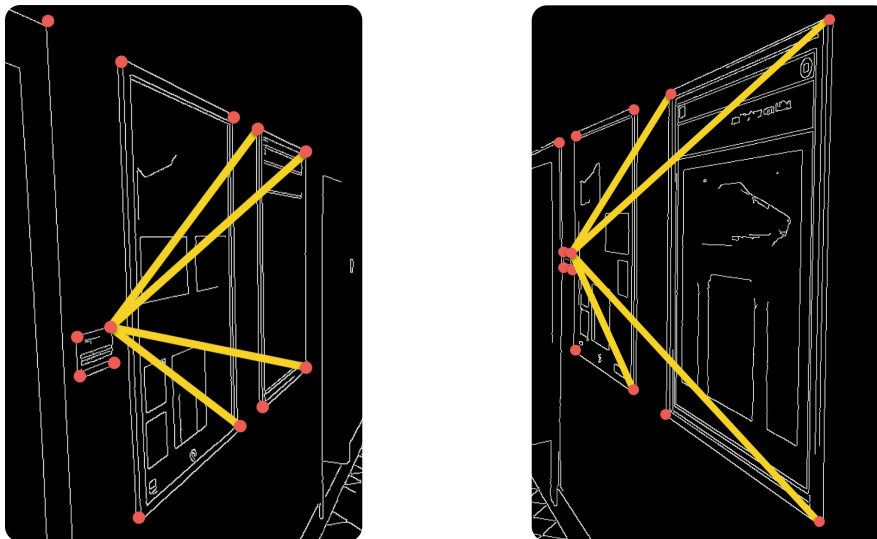


Abbildung 5.1: Voraussetzungen für die Plausibilitätstests: Es müssen vier Nachbarpunkte über die virtuellen Linienmerkmale bestimmt worden sein wie in Kapitel 4.3

5.2.2 Plausibilitätstests

Der erste Test überprüft die Anzahl der Punkte auf deren konvexen Hülle. Wenn beide Fünferpaare in beiden Bildern aus denselben Punkten bestehen, so stimmt auch die Anzahl der Punkte auf der konvexen Hülle in beiden Bildern überein (siehe Kapitel 3.1.2). Daher wird jeweils die Konvexe Hülle der fünf Punkte bestimmt und die Anzahl der Punkte, die diese Hülle aufspannen, verglichen. Sollte die Anzahl der

Punkte auf der komplexen Hülle in beiden Bildern übereinstimmen, wird danach die Reihenfolge der Nachbarpunkte bestimmt. Dass die Topologie stimmt, wurde bereits getestet und ist Vorbedingung. Um im Folgenden die P^2 -Invarianten I_1 und I_2 berechnen zu können, ist deren identische Festlegung in beiden Bildern notwendig. Wie bei der Überprüfung der Nachbarschaftsreihenfolge wird nun auch für deren eindeutige Bestimmung jeweils der Richtungswinkel vom untersuchten Punkt aus zu den vier Nachbarpunkten bestimmt. Anschließend werden die vier Nachbarpunkte nach der Größe der Richtungswinkel geordnet. Danach folgt die Überprüfung der P^2 -Invarianten I_1 und I_2 , die in beiden Bildern gleich groß sind, wenn die fünf Punkte dieselben sind und in einer Ebene liegen (vergleiche Abbildungen 3.5 und 5.2). Für jedes untersuchte Punktepaar wird dabei gezählt, wie viele P^2 -Invarianten identisch sind und die Anzahl an Übereinstimmungen wird als Indikator für die Bestätigung der Kandidaten in einer Liste gespeichert.

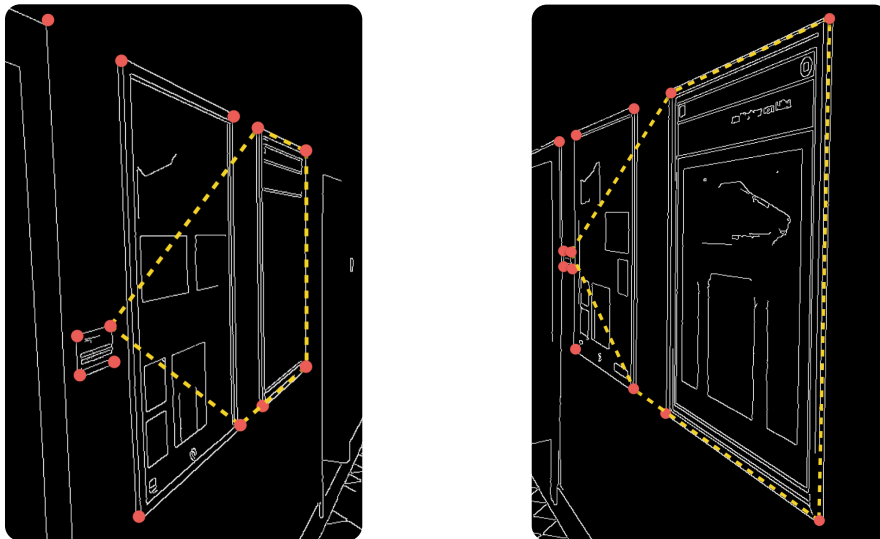


Abbildung 5.2: Plausibilitätstests bezüglich konvexer Hülle und P^2 -Invarianten mit den gleichen fünf Punkten durchführen. Die Anzahl der Punkte auf der Konvexen Hülle, die Reihenfolge der Punkte, sowie die Invariante I_1 und I_2 müssen übereinstimmen, wenn die fünf Punkte die gleichen sind.

5.2.3 Logische Reduzierung der Kandidatenliste

Die Liste der Kandidaten und ihrer Anzahl an Bestätigungen durch die P^2 -Invarianten wird danach anhand ihrer Bestätigungszahlen, die aus den Plausibilitätstests erstellt wurden, logisch reduziert. Dafür wird die Liste zunächst anhand der Bestätigungszahlen geordnet. Das Kandidatenpaar mit der höchsten Bestätigung wird als richtig angenommen und alle anderen Paare, welche die gleichen Koordinaten enthalten,

werden entfernt. Diese Prozedur wird mit allen weiteren Einträgen der Liste durchgeführt, bis nur noch eindeutige Paare vorhanden sind. Die übrig gebliebenen Paare sind somit diejenigen, deren Nachbarn, welche durch identische Linienfeatures prognostiziert wurden, am häufigsten durch die Plausibilitätstests bestätigt wurden. Sollten richtige Paare in dieser Liste enthalten sein, so können diese und die anderen korrekten Paare mit dem folgenden abschließenden Test herausgefiltert werden.

5.2.4 Abschließende Überprüfung durch virtuelle Linienfeatures

In diesem abschließenden Test werden alle durch die Plausibilitätstests (Kapitel 5.2.2) gefundenen Kandidatenpaare gegen die im vorherigen Schritt (Kapitel 5.2.3) aufgestellten Punktkorrespondenzen getestet, von denen angenommen wird, dass zumindest einige davon richtig sind. Sollte eine Punktkorrespondenz der Liste mit den maximalen Bestätigungen tatsächlich richtig sein, so muss auch das virtuelle Linienfeature zu einem anderen richtigen Kandidatenpaar in beiden Bildern identisch sein, wenn sich beide auf der gleichen projektiven Ebene befinden. Daher werden nun je Kandidatenpaar, welches durch die Plausibilitätstests positiv bestätigt wurde (Kapitel 5.2.2), alle Linienfeatures zu den Koordinaten der als richtig vermuteten Korrespondenzen (Kapitel 5.2.3) verglichen. Der Vergleich der Linienfeatures erfolgt wie in Kapitel 4.3.3 beschrieben. Für jedes Kandidatenpaar wird die Anzahl der bestätigten Linienfeatures zu den als richtig angenommenen Punkten gespeichert. Unabhängig davon, ob ein Kandidatenpaar in der Liste der vermuteten richtigen Zuordnungen enthalten war, können anschließend anhand der Anzahl der Bestätigungen die richtigen Punktkorrespondenzen gefunden werden (siehe Abbildung 5.3). Dies ist theoretisch sogar dann möglich, wenn nur eine einzige der angenommenen richtigen Zuordnungen tatsächlich korrekt war, da die falschen Annahmen keine oder deutlich weniger identische Linienverhältnisse aufweisen als die richtigen.

Die so gefundenen bestätigten Punktkorrespondenzen werden nachfolgend wie oben beschrieben erneut logisch reduziert und die Punkte mit der maximalen Anzahl an bestätigten Linienfeatures werden als endgültiges Ergebnis ausgegeben.

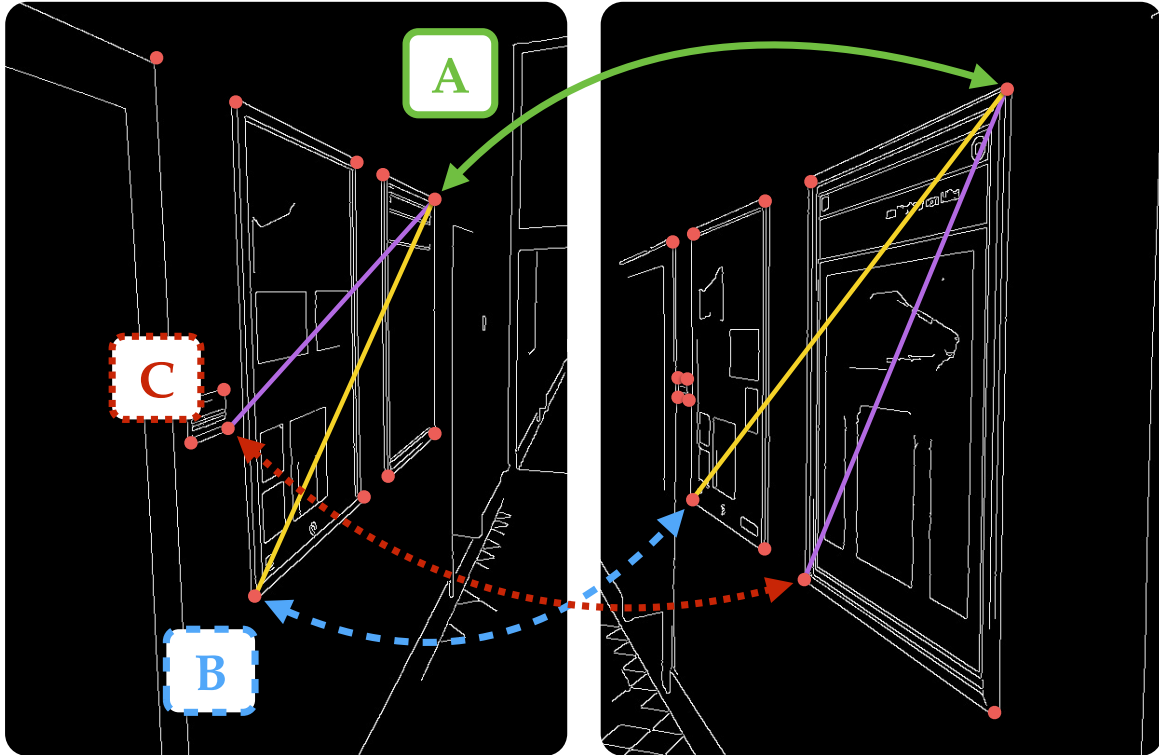


Abbildung 5.3: Abschließende Überprüfung: Punktkorrespondenz A (grün dargestellt) ist Teil der Punktkorrespondenzen, die nach den Plausibilitätstests die höchste Anzahl an Bestätigungen haben. Von diesen wird vermutet, dass darin zumindest eine Korrespondenz enthalten ist, welche richtig ist. Punktkorrespondenz B (blau gestrichelt) wird überprüft, indem die Linienfeatures in beiden Bildern, die von B nach A gehen (gelb dargestellt), übereinstimmen. Ist dies wie im Bild der Fall, wird die Punktkorrespondenz als richtig bewertet. Punktkorrespondenz C ist ein Beispiel für eine falsche Korrespondenz (rot gepunktet). Die Linienfeatures von C nach A (violett dargestellt) sind nicht identisch, weshalb die Punktkorrespondenz keine Bestätigung erhält. Nacheinander werden alle anderen Kandidaten, die bei den Plausibilitätstests ein positives Ergebnis hatten, gegen alle als richtig eingestuften Korrespondenzen (z.B. Punktkorrespondenz A) geprüft.

5 Neues Konzept: kombinierte Linien- und Flächenmerkmale

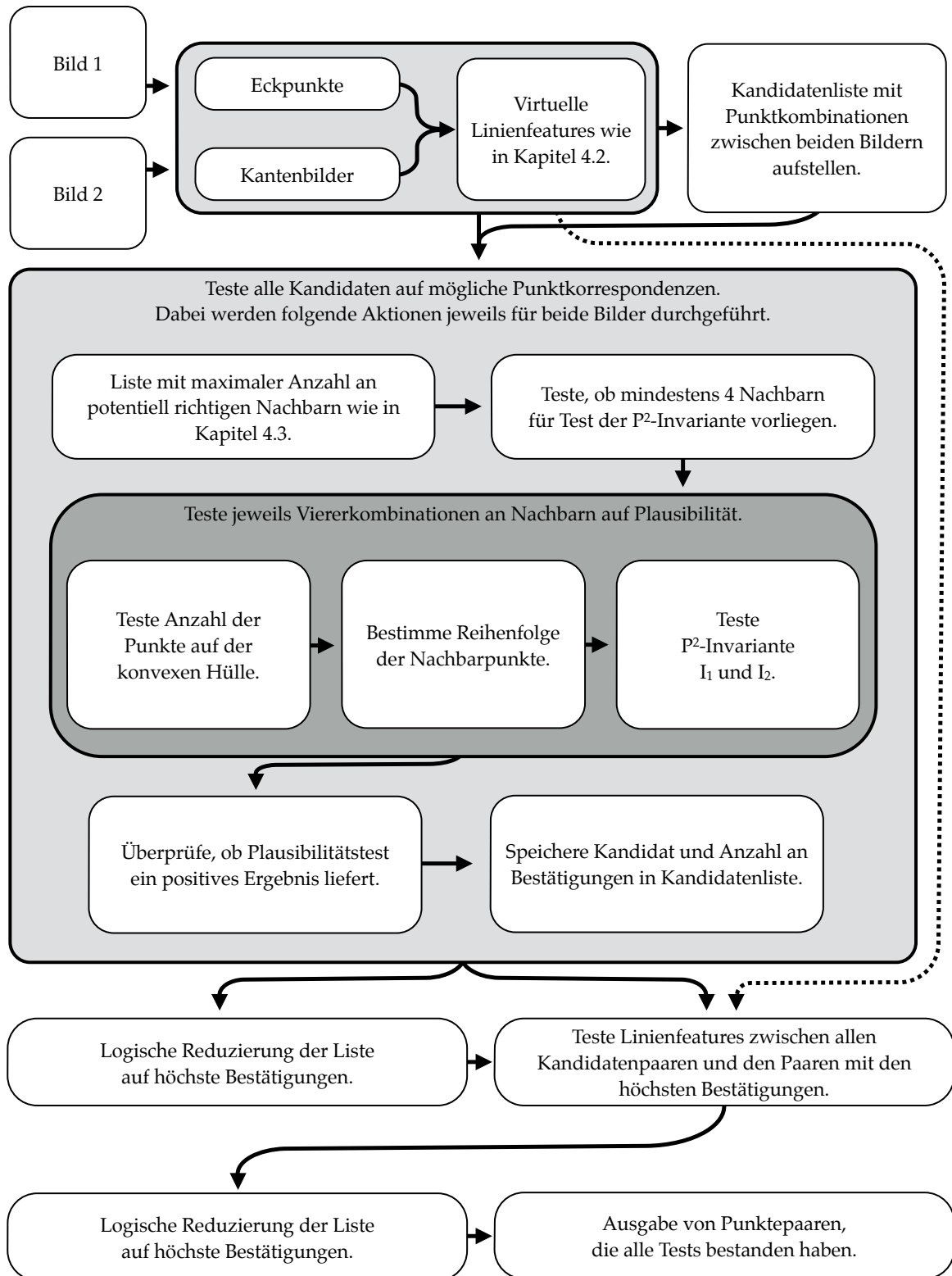


Abbildung 5.4: Das neue Konzept der Kombination aus Linien- und Flächenmerkmalen.

6 Evaluation

In diesem Kapitel wird der in dieser Arbeit vorgestellte neue Algorithmus zunächst anhand von synthetischen Daten evaluiert (Kapitel 6.2). Danach wird der Algorithmus in Kapitel 6.3 an Realdaten getestet. Im Anschluss daran werden Vergleiche mit bestehenden Verfahren in Kapitel 6.5 besprochen.

6.1 Synthetische Testdaten

Ziel der Evaluation ist es, die prinzipielle Anwendbarkeit und die Stärken und Schwächen in Abhängigkeit der jeweiligen Gegebenheiten zu zeigen. Die Anforderungen an das Verfahren, die in Kapitel 1.2 aufgestellt wurden, sind zusammengefasst folgende:

- Besonders große Blickwinkelunterschiede sollen handhabbar sein.
- Nahezu beliebige perspektivische Deformationen, die besonders stark bei sehr kleinen Kamerakonstanten auftreten, sollen handhabbar sein.
- Das Verfahren soll robust gegenüber radiometrischen Veränderungen im Bild sein.
- Das Konzept soll auf möglichst unterschiedliche Bilder übertragbar sein.
- Teilausschnitte von Bildern sollen handhabbar sein.
- Verdeckungen im Bild sollen handhabbar sein.
- Der Algorithmus soll wenige einfache Rechenoperationen benutzen.
- Der Algorithmus soll gegenüber vorhandenen projektiv invarianten Verfahren eine deutliche Reduktion der Rechenzeit erreichen.

Um zu überprüfen, ob diese Anforderungen erfüllt sind, werden synthetische Testdaten erstellt, die diese Herausforderungen kontrolliert wiedergeben. Um unterschiedliche Blickwinkel zu simulieren, wird ein Ausgangsbild (Abbildung 6.2 (a)) mit Hilfe einer projektiven Transformation perspektivisch abgebildet. Dies geschieht in Schritten von 5° , wobei jeweils gleichzeitig eine Rotation (Winkel κ) und eine Neigung (Winkel ω) des Bildes durchgeführt wird. Da bei einer Abbildung mit einer Kamera die verwendete Kamerakonstante einen erheblichen Einfluss auf die Abbildung

und deren perspektivische Deformation hat, wird diese in Form eines Faktors P mit berücksichtigt. P ergibt sich dabei wie in Abbildung 6.1 dargestellt. Beträgt der Fak-

$$P = \frac{\text{Bilddiagonale}[\text{Pixel}]}{\text{Kamerakonstante}[\text{Pixel}]} \quad (6.1)$$

Abbildung 6.1: Faktor P für die perspektivische Deformation.

tor $P = 1$, so kommt dies dem menschlichen Sehen sehr nahe. Die Einflüsse von unterschiedlichen P werden in Kapitel 6.2.3 behandelt. Abbildung 6.2 zeigt die synthetischen Testdaten auszugsweise.

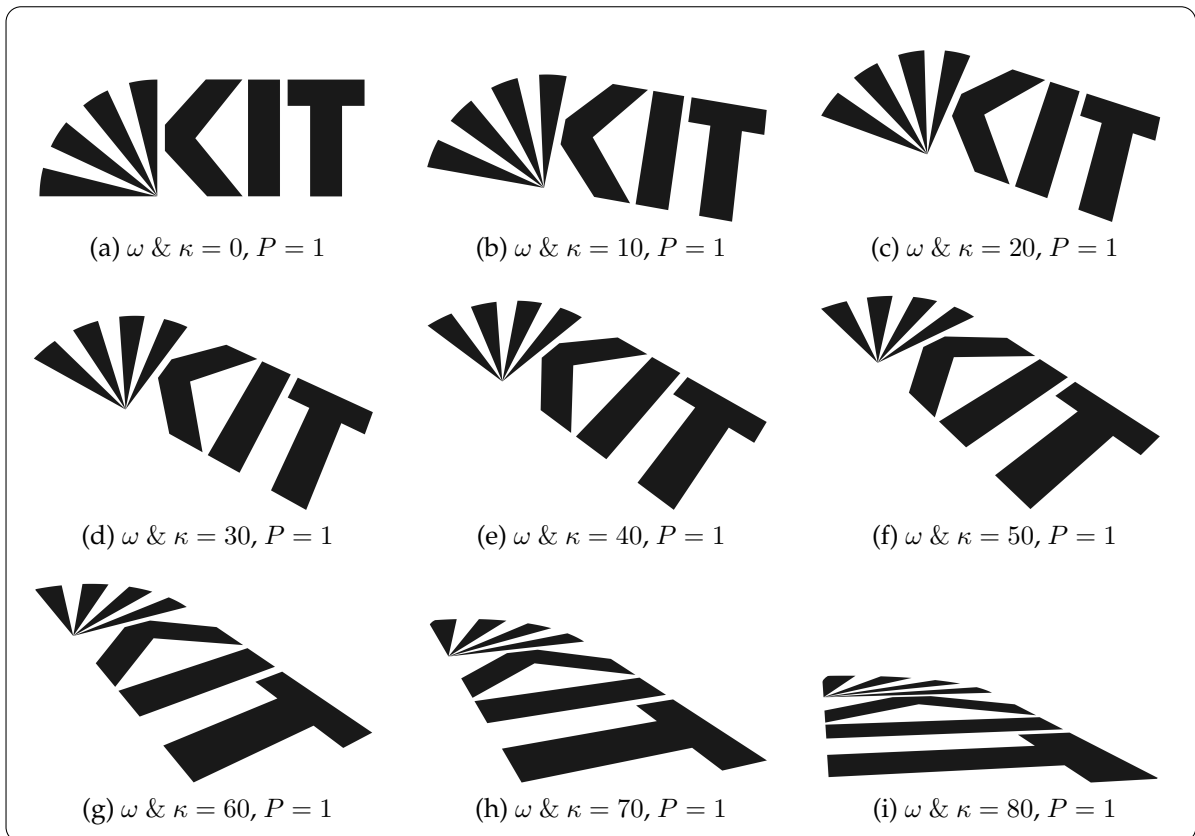


Abbildung 6.2: Synthetisch produzierte perspektivische Projektionen des Ursprungsbildes (a) unter Benutzung der Drehwinkel ω und κ . Auszugsweise ist jeder zehnte Schritt dargestellt.

Diese Testdaten werden im folgenden Kapitel verwendet, um beispielsweise den Einfluss unterschiedlicher Blickwinkel auf die produzierten Ergebnisse zu untersuchen.

6.2 Sensitivitätsanalysen

6.2.1 Eckpunkt-Features und Blickwinkelveränderungen

Der Algorithmus benutzt automatisch detektierte Eckpunkte in beiden Stereobildern und ordnet diese wie zuvor beschrieben einander zu. Wie in Kapitel 3.2.3 angesprochen, gibt es viele unterschiedliche Eckpunkt-Features, die dafür verwendet werden können. Zwei davon werden hier gesondert betrachtet, um zu zeigen, dass die Wahl der Eckpunktfeatures einen Einfluss auf das Ergebnis haben kann. Die Anzahl der Eckpunkte eines Harris-Operators [Harris & Stephens, 1988] und der Eckpunkte eines Förstner-Operators [Förstner & Gülch, 1987] oder des neueren SFOP-Operators [Förstner *et al.*, 2009], bei dem Kreuzungspunkte von Linien im Bild verwendet werden, sind unterschiedlich und deren Positionen sind teilweise an unterschiedlichen Bildkoordinaten.

Ergebnisse der beiden unterschiedlichen Eckpunktdetektoren sind in Abbildung 6.3 dargestellt. Eckpunkte einer Methode als besser als die einer anderen Methode darzu-

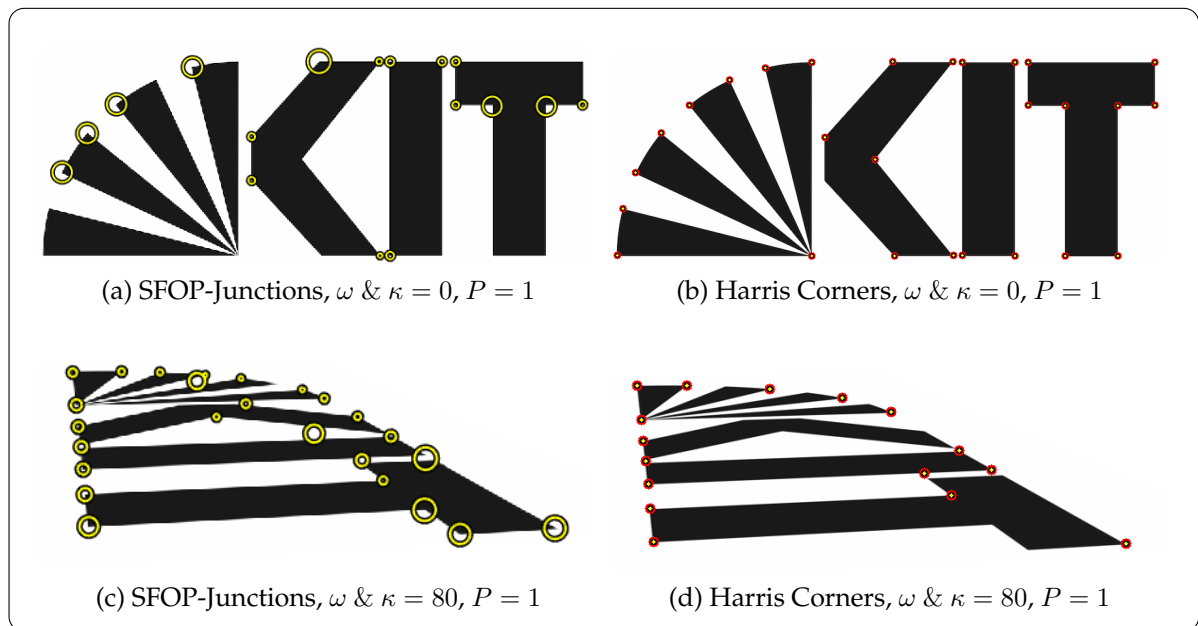


Abbildung 6.3: Unterschiedliche Anzahl und Position an Eckpunkten je nach Detektor.

stellen ist schwierig. Beide haben ihre Vor- und Nachteile. So können die Positionen von SFOP-Junctions oftmals zuverlässiger in beiden Bildern wiedergefunden werden, da deren Positionen mit Hilfe einer größeren Nachbarschaft um den Eckpunkt herum und durch ein Linienschnittverfahren subpixelgenau errechnet wurden. Gerade bei perspektivischen Abbildungen hat dies den Vorteil, dass sich diese Positionen

zwischen zwei Blickwinkeln weniger verändern als beispielsweise die von Harris-Eckpunkten. Ein Nachteil ist, dass nur dort Eckpunkte erkannt werden, wo auch in der berücksichtigten Nachbarschaft Linienschnitte detektiert wurden. Dies spiegelt sich auch in der Anzahl der gefundenen Punkte wieder, die bei SFOP-Junctions in der Praxis oft geringer ist als bei Harris-Eckpunkten, wenn die gleichen Bilder verglichen werden (siehe Abbildung 6.4).

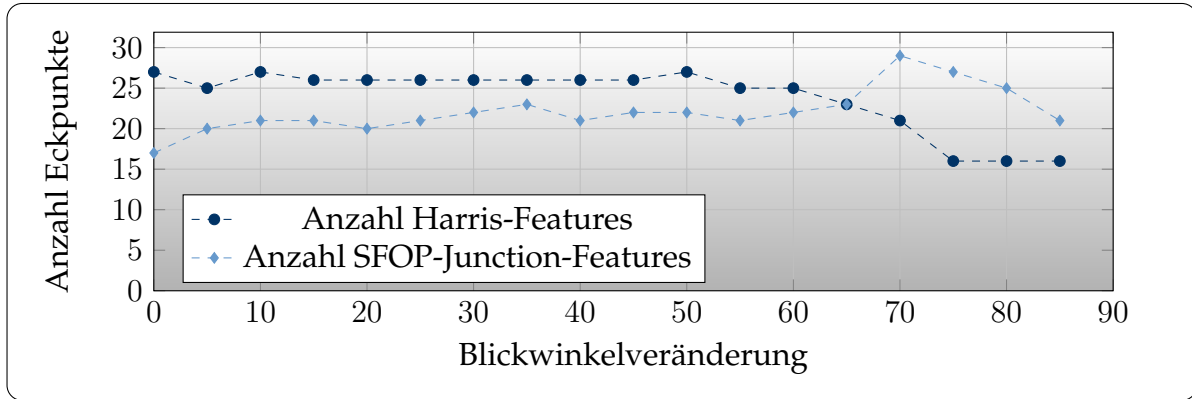


Abbildung 6.4: Anzahl gefundener Eckpunkte in Abhängigkeit vom Neigungswinkel.

6.2.2 P^2 -Invariante I_1 und I_2 und Blickwinkelveränderungen

In diesem Unterkapitel wird das Verfahren dahingehend getestet, ob es Bildinhalte, welche aus unterschiedlichen Perspektiven aufgenommen wurden, einander zuordnen kann. Dazu werden die erstellten Testdaten benutzt, die im Kapitel 6.1 beschrieben wurden. Es werden jeweils die geeigneten Bilder und das untransformierte Ursprungsbild für einen Zuordnungstest verwendet. Die Ergebnisse der Punktzuordnung können wie in Abbildung 6.5 visualisiert aussehen. Zu sehen ist jeweils oben das untransformierte Bild und darunter das perspektivisch transformierte. Jede grüne Linie gibt eine vollständige bzw. eindeutige projektive Transformation an, da die Korrespondenz auf mindestens fünf Punkten beruht. Wie aus der Abbildung hervorgeht, können selbst bei einer sehr starken Blickwinkelveränderung von 85° noch zwei richtige Punktkorrespondenzen errechnet werden. Dabei ist hervorzuheben, dass sowohl für die Erstellung dieses Ergebnisses, als auch bei der Berechnung aller anderen Ergebnisse, keine statistischen Ausreißertests, wie beispielsweise RANSAC [Fischler & Bolles, 1981], notwendig waren.

Dieser Versuch wurde für alle Blickwinkelveränderungen der Winkel ω und κ zwischen den Werten 0° und 85° in Schritten von 5° durchgeführt. Dabei wurden jeweils zusätzlich zu den Blickwinkeln auch die Spannweiten der zugelassenen Abweichungen der P^2 -Invarianten I_1 und I_2 verändert, um deren Einfluss auf die Endergebnisse

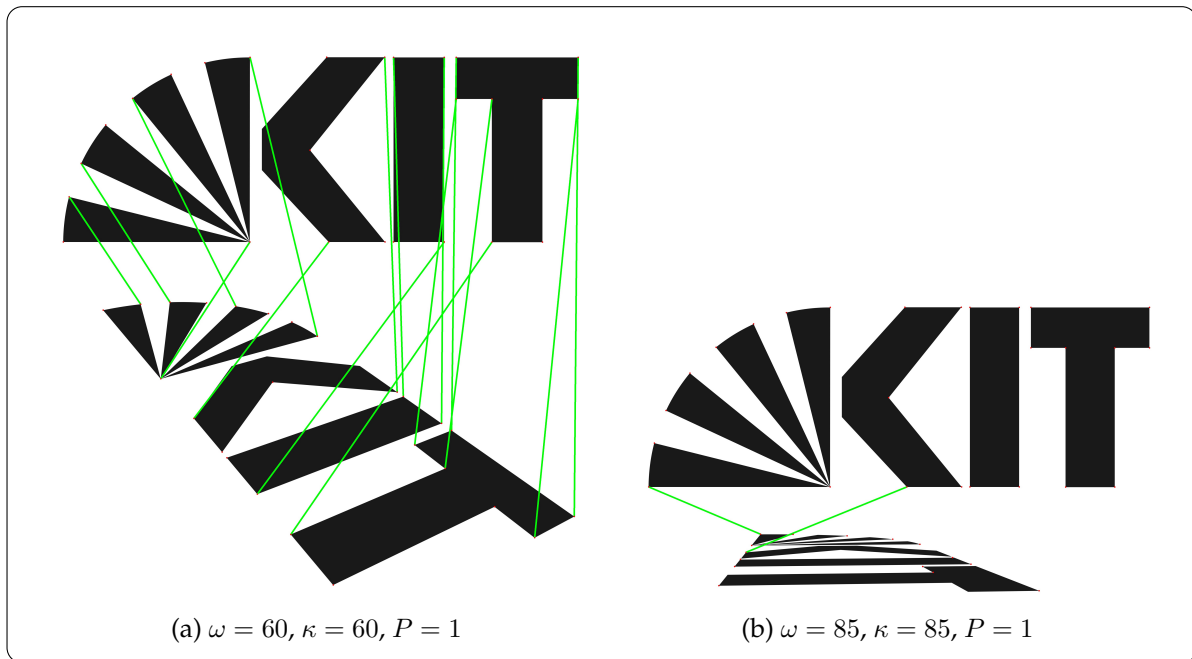


Abbildung 6.5: Ergebnisse der Punktzuordnung 60 und 85 Grad. Hier wurden Harris-Eckpunkte verwendet. Jede grüne Linie entspricht dabei einer vollständigen bzw. eindeutigen projektiven Transformation, da die Korrespondenz auf mindestens fünf Punkten beruht. Hervorzuheben ist, dass für dieses Ergebnis, wie auch für alle anderen, keine statistischen Ausreißertests notwendig waren.

zu untersuchen. Größere zugelassene Abweichungen bei den Plausibilitätstests führen dazu, dass mehr Kandidaten für Punktkorrespondenzen erstellt werden. Diese größere Anzahl an Kandidaten enthält dann jedoch weniger richtige Zuordnungen, was anschließend verursacht, dass bei der abschließenden Überprüfung der Punkte (vergleiche Kapitel 5.2.4) falsche Zuordnungen als richtig angenommen werden können. Ein Grund, warum überhaupt größere Abweichungen bei den P^2 -Invarianten zugelassen werden müssen ist, dass bei der Verwendung von Realdaten die Verzerrung des Bildes mit berücksichtigt werden muss, weshalb die Kenngrößen I_1 und I_2 niemals exakt übereinstimmen. Hier muss beim Vergleich die höhere Spannweite der Abweichungen zugelassen werden, um korrekte Ergebnisse zu erzielen. Das Ergebnis dieser Versuchsreihe ist in den Abbildungen 6.6 und 6.7 dargestellt. Zu sehen sind vier Datenreihen. Jede davon steht für eine andere zugelassene Spannweite bei Vergleich der P^2 -Invarianten in den Plausibilitätstests. Die Spannweiten wurden bei Werten von 0.05, 0.1, 0.2 und 0.3 getestet.

Die Ergebnisse in Abbildungen 6.6 zeigen, dass wie erwartet für höhere zugelassene Abweichungen mehr Punktkorrespondenzen gefunden wurden. Zu beachten ist, dass selbst bei sehr starken Blickwinkelveränderungen (70° bis 85°) immer noch

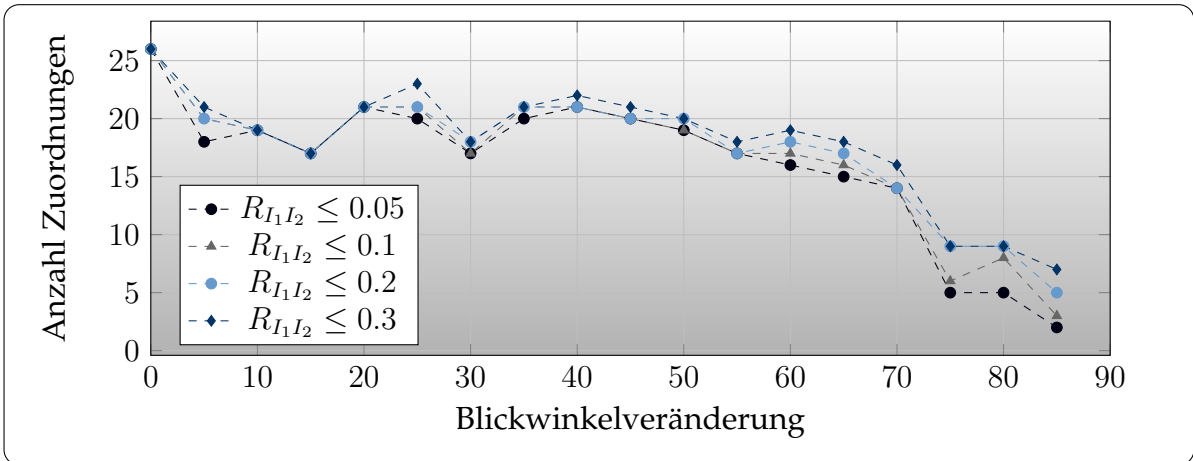


Abbildung 6.6: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.

Punktzuordnungen erkannt werden. Die Richtigkeit dieser Punktzuordnungen ist in Abbildung 6.7 zusammengefasst. Bei kleinen Veränderungen des Blickwinkels bis 35° gegenüber dem Originalbild zeigen sich keine Unterschiede zwischen den gewählten zugelassenen Spannweiten. Alle Punktzuordnungen sind in diesem Bereich richtig. Ab 40° ist bei der höchsten zugelassenen Abweichung von 0.3 zum ersten Mal ein Ausreißer enthalten. Bei den Spannweiten 0.1 und 0.2 ist der erste Ausreißer bei 45° , bei 0.05 erst bei 55° . Aus den Kurven ist der erwartete Zusammenhang zwischen einer niedrigeren erlaubten Abweichung und mehr richtigen Punktzuordnungen direkt ersichtlich. Trotz einiger Ausreißer ist aber bei allen Spannweiten die Anzahl der richtigen Zuordnungen deutlich höher als die der falschen.

Ebenfalls interessant ist die Rechenzeit, die das Programm benötigte, um die Punktzuordnungen zu erstellen. Die geringste zugelassene Spannweite resultierte auch in der geringsten Rechenzeit von durchschnittlich 16 Sekunden pro Bildpaar. Die größeren Spannweiten verursachten höhere Rechenzeiten. Dies ist damit begründet, dass bei größeren Spannweiten natürlich mehr Punkte als Kandidaten für Korrespondenzen vorgeschlagen werden, die dann alle getestet werden müssen. In den durchgeführten Untersuchungen hat sich die Rechenzeit ab einer Spannweite von 0.1 jedoch nicht mehr wesentlich verändert. Die Ergebnisse sind in Abbildung 6.8 dargestellt.

6.2.3 Unterschiedliche perspektivische Deformationen

Um zu demonstrieren, dass auch starke perspektivische Deformationen vom Algorithmus richtig erkannt werden, wurden neben den Testreihen mit normalen perspektivischen Deformationen auch deutlich kürzere Kamerakonstanten simuliert. Kurze Kamerakonstanten führen im Bild zu extremen perspektivischen Verkürzungen.

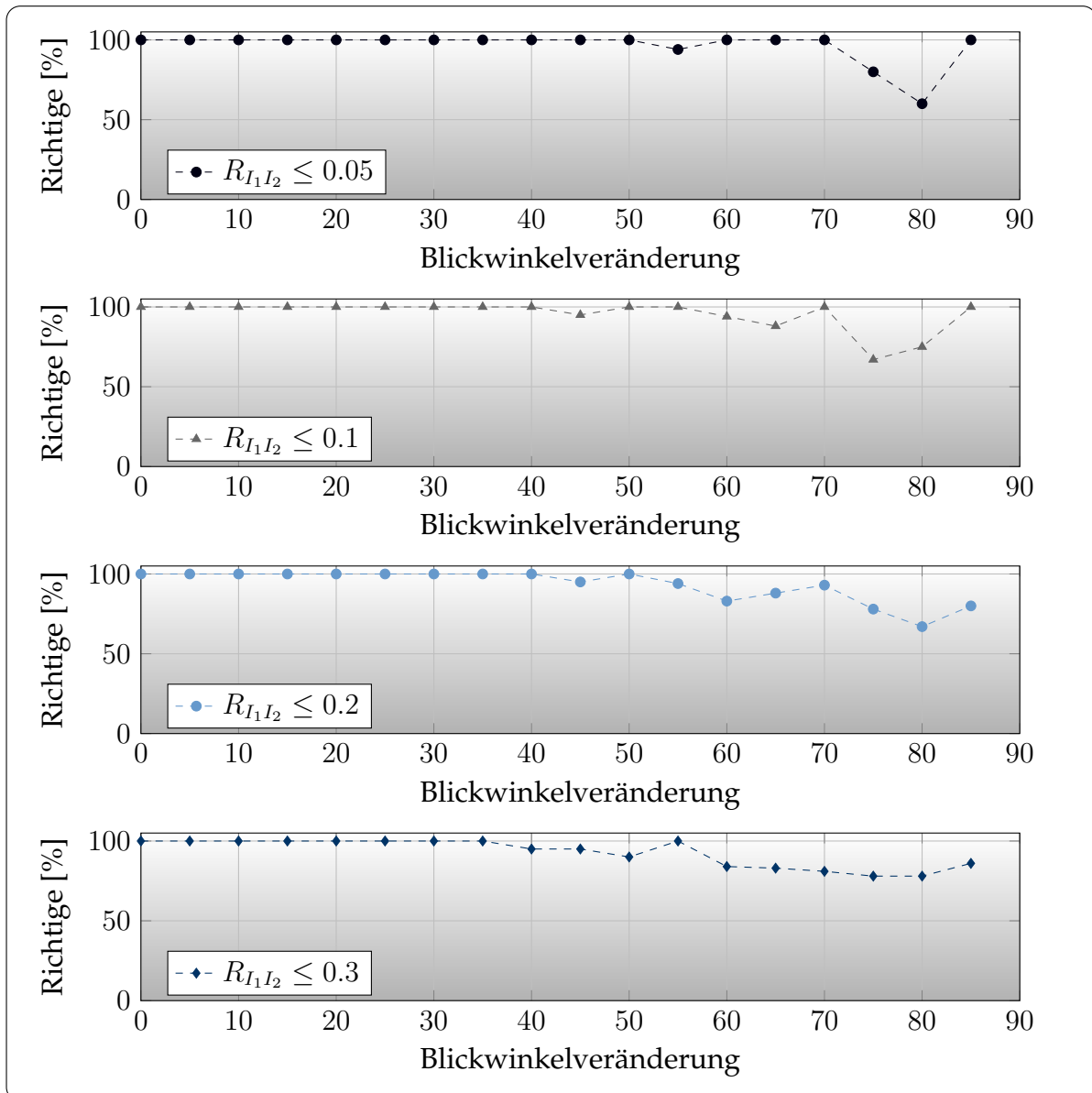


Abbildung 6.7: Richtige Zuordnungen bei unterschiedlichen zugelassenen Spannweiten in Prozent. Mit größeren zugelassenen Spannweiten erhöht sich die Unsicherheit bei starken Blickwinkelveränderungen. Dass die Kurven am Ende bei 85° nochmals ansteigen ist auf die numerische Unsicherheit, bedingt durch die geringe Anzahl an gefundenen Zuordnungen, zurückzuführen.

Abbildung 6.9 zeigt drei unterschiedliche simulierte perspektivische Deformationen. Die erste Abbildung 6.9 (a) zeigt zum Vergleich eine rein affine Abbildung, hier ist $P \approx 0$. Abbildung 6.9 (b) zeigt eine normale, dem menschlichen Sehen nachempfundene Perspektive mit $P = 1$ und Abbildung 6.9 (c) zeigt eine Abbildung mit einer kurzen Kamerakonstante mit $P = 1.5$. Der Unterschied zwischen der affinen Abbildung

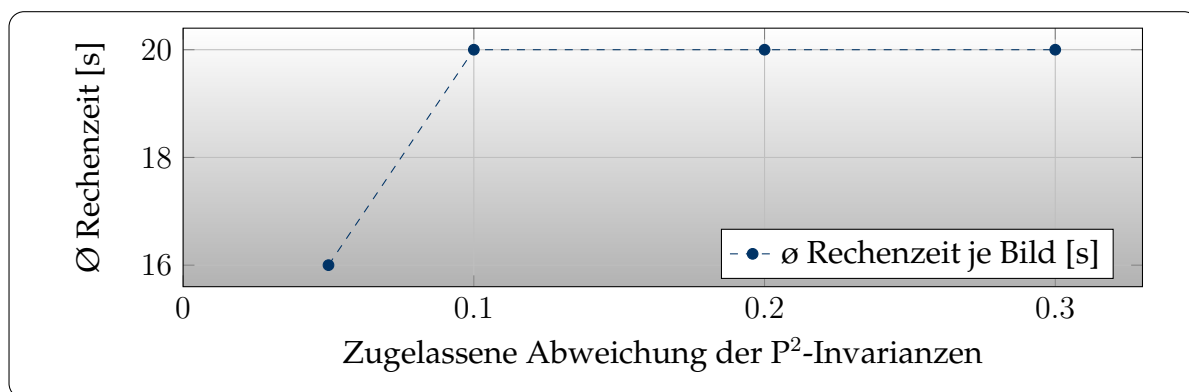


Abbildung 6.8: Laufzeiten bei unterschiedlichen zugelassenen Spannweiten.

und der perspektivischen ist sofort sichtbar. In der affinen Abbildung bleiben parallele Geraden parallel, in der perspektivischen Abbildung schneiden sich ursprünglich parallele Geraden in einem Fluchtpunkt. Auch die perspektivische Verkürzung ist gut sichtbar. So ist das K von KIT in Abbildung 6.9 (c) deutlich kleiner und enger zulaufend als in den Abbildungen daneben.

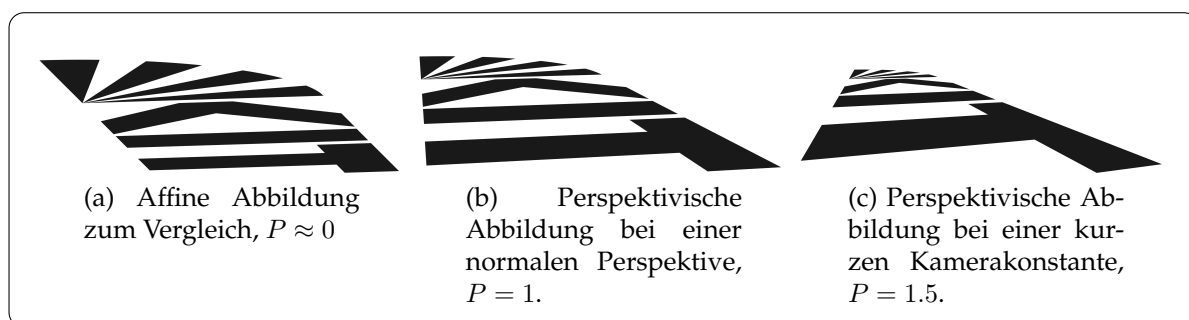


Abbildung 6.9: Simulation unterschiedlicher Kamerakonstanten. Die Rotationen ω und κ betragen jeweils 80 Grad. (a) Zeigt eine affine Abbildung ohne die perspektivischen Effekte. (b) zeigt eine perspektivische Abbildung mit einer simulierten normalen Perspektive. (c) zeigt eine perspektivische Abbildung mit einer simulierten kurzen Kamerakonstante, hier sind die perspektivischen Verkürzungen besonders gut sichtbar.

Abbildung 6.10 visualisiert die Ergebnisse der Punktzuordnungen. Grüne Linien zeigen errechnete Punktkorrespondenzen zwischen beiden Bildern an. Zu sehen sind zwei unterschiedliche perspektivische Transformationen des Ursprungsbildes, die jeweils mit dem untransformierten Bild verglichen wurden. Die perspektivische Deformation ist hier besonders stark mit einem Faktor $P = 1.6$.

Für diese Untersuchung wurden verschiedene P von 0 bis 1.6 benutzt. Die Ergebnisse sind in Abbildung 6.11 dargestellt. Abgebildet sind jeweils die Anzahl der gefunde-

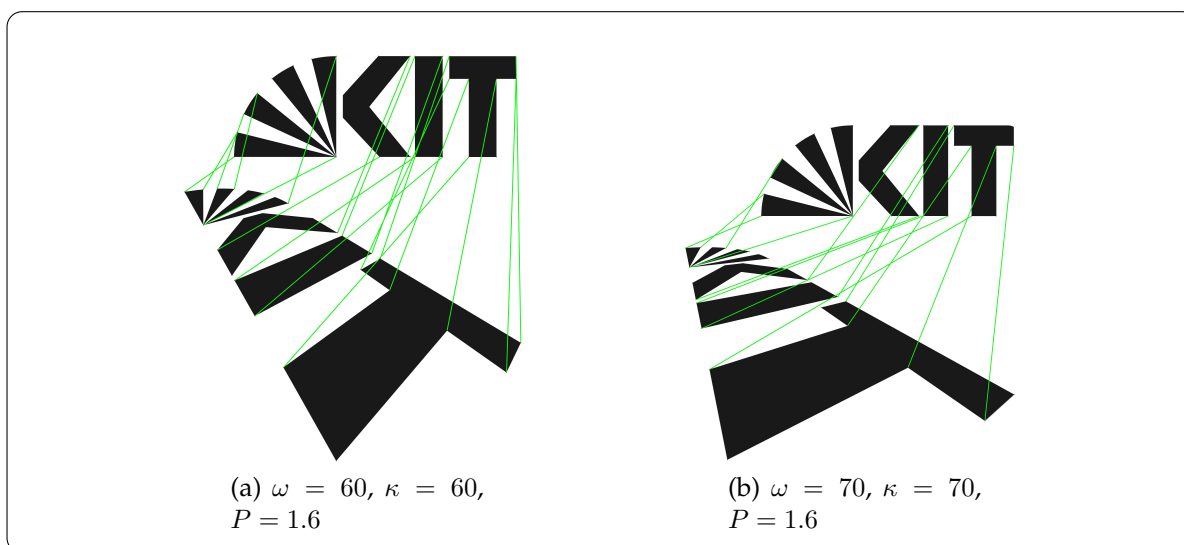


Abbildung 6.10: Ergebnisse 60 und 70 Grad bei einer simulierten kurzen Kamerakonstante, was sich im transformierten Bild in einer stärkeren perspektivischen Verkürzung äußert.

nen Übereinstimmungen in Abhängigkeit vom gewählten Faktor P sowie der Anteil der davon richtigen Zuordnungen in Prozent. Aus den Untersuchungen geht hervor, dass der Einfluss der perspektivischen Deformation auf das Endergebnis der Punktzuordnungen keinen erkennbaren Einfluss hat. Bei der Blickwinkeländerung von 80° ist zwar ein Rückgang der richtigen Zuordnungen zu sehen, allerdings ist dies auch der Bereich, in dem durch die extreme Veränderung des Blickwinkels deutlich weniger Zuordnungen gemacht wurden. Insgesamt zeigt sich keine deutliche Veränderung der Ergebnisse bei unterschiedlichen simulierten Kamerakonstanten, was auch zu erwarten war, da alle benutzen Funktionen des Algorithmus perspektivisch invariant sind.

6.2.4 Exemplarischer Test von unterschiedlichen Radiometrien

Bevor in Kapitel 6.3 in Abbildung 6.26 (b) die radiometrische Invarianz erfolgreich anhand von Realdaten demonstriert wird, wird hier nochmals ein exemplarischer Test durchgeführt und dessen Zwischenschritte visualisiert. Der Algorithmus ist deshalb radiometrisch invariant, weil als Eingangsdaten lediglich Eckpunkte und Kantenbilder verwendet werden. Wenn die Eckpunkte und die Kanten im Bild durch geometrische Strukturen oder unterschiedliche Materialien verursacht wurden, so sind diese meist sehr zuverlässig in anderen Bildern erkennbar. Daher haben unterschiedliche Intensitätsinformationen keinen oder nur einen sehr geringen Einfluss auf das Ergebnis. In Abbildung 6.12 ist die erfolgreiche Punktzuordnung zweier geeigneter Bilder dargestellt, die zudem unterschiedliche Intensitäten in den einzelnen

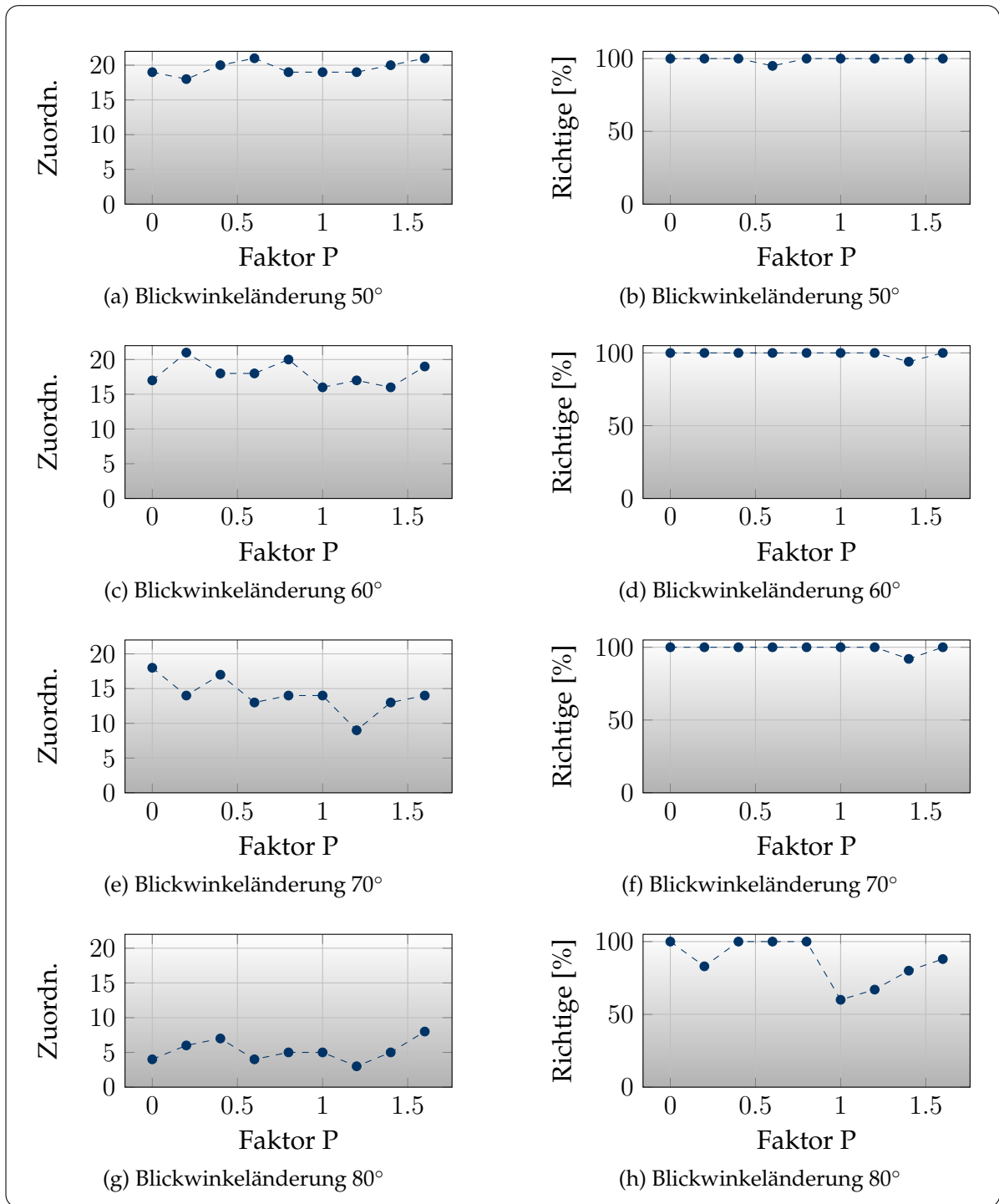


Abbildung 6.11: Die Auswirkungen von unterschiedlichen perspektivischen Deformationen bedingt durch unterschiedliche Kamerakonstanten.

Kanälen aufweisen. Das Grün im einen Bild wird zu Schwarz im anderen. Trotzdem

bleiben die meisten Eckpunkte auch im anderen Bild Eckpunkte und die Kanten bleiben Kanten. Die Eckpunkt Algorithmen, sowie die Kantendetektoren arbeiten in erster Linie mit Einkanalbildern. Bei RGB-Mehrkanalbildern wird dazu zunächst das Bild in den HSV-Farbraum (englisch für hue=Farbwert, saturation=Sättigung und value=Hellwert) transformiert. Die Helligkeitswerte (V) werden danach mit einem Bilateralfilter geglättet, bevor sie als Eingangsdaten für die Eckpunkt- und Kantendetektion verwendet werden.

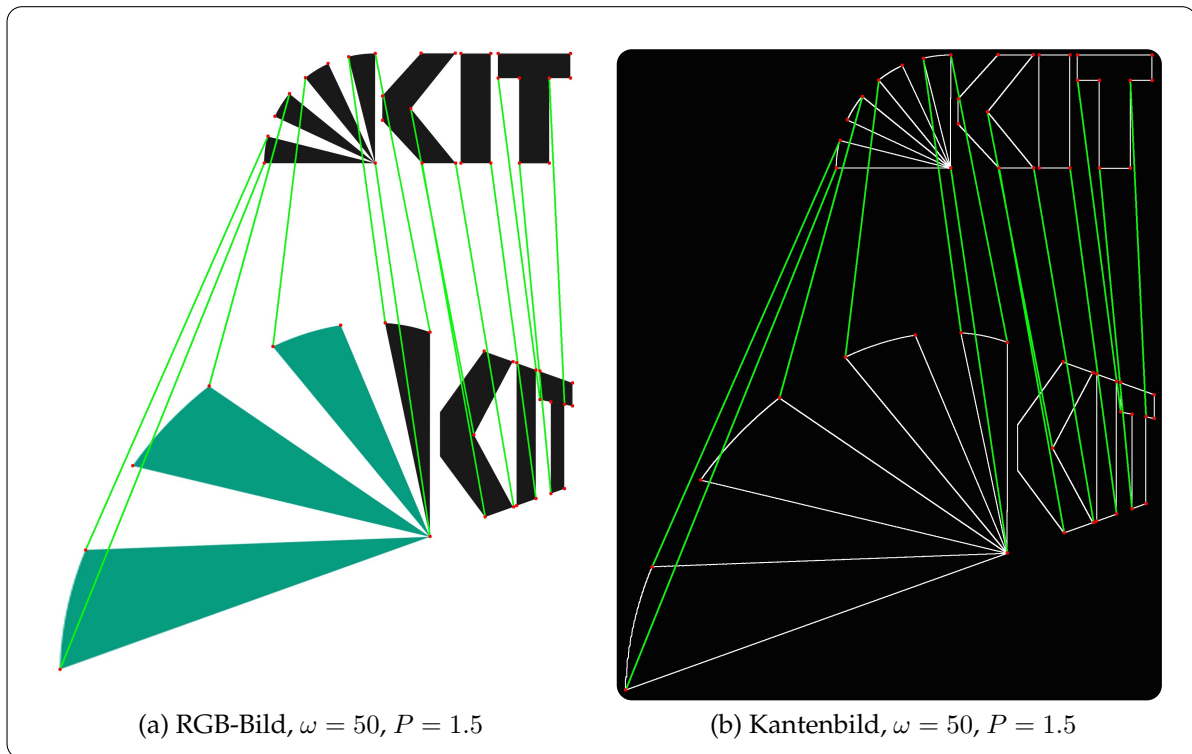


Abbildung 6.12: Unterschiedliche Radiometrien haben keinen Einfluss auf die zugrunde liegende Struktur. $\omega = 50^\circ$, $P = 1.5$

6.2.5 Punktkorrespondenzen bei Untermengen

Bisher wurden Bilder miteinander verglichen, die in beiden Fällen die gleiche Anzahl an Eckpunkten hatten. In diesem Unterkapitel wird untersucht, wie sich fehlende Information auswirkt. Dies geschieht, wenn eines der Bilder zwar die gleichen Eckpunkte enthält, jedoch nicht alle identisch sind. Dafür wurden Bilder erstellt, wie sie in Abbildung 6.13 zu sehen sind. Hier fehlen Bildinformationen im Ursprungsbild.

Getestet wurde, ob der Algorithmus auch hier Punktzuordnungen findet und ob diese richtig sind. Die Ergebnisse sind in den Abbildungen 6.14 und 6.15 dargestellt. Unabhängig vom Blickwinkel zeigt sich, dass die Anzahl an Zuordnungen geringer

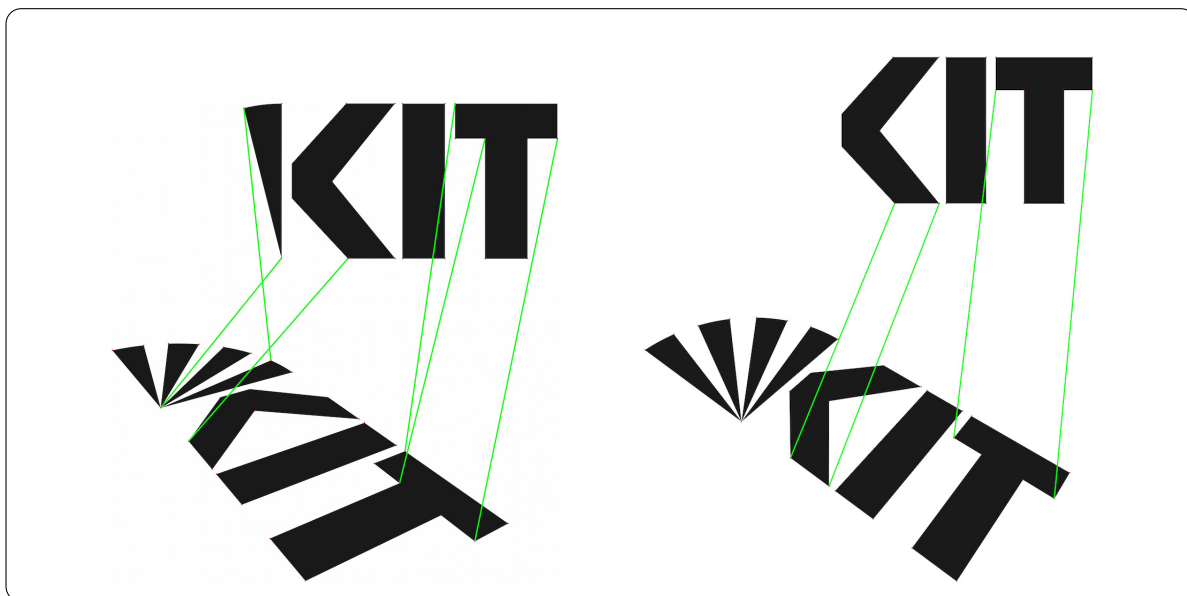


Abbildung 6.13: Unterschiedliche Teilbereiche, die jeweils Teilmengen des anderen Bildes darstellen.

wird, je kleiner die Teilmenge gewählt wird. Das liegt daran, dass zum einen weniger Punkte vorhanden sind, die zugeordnet werden können, und zum anderen, dass bedingt durch weniger Kanteninformationen auch weniger Punkte zugeordnet werden. Bei kleinen Teilmengen sind zu wenig und gegebenenfalls mehrdeutige Informationen vorhanden, um überhaupt mit den Plausibilitätstests beginnen zu können. Hierfür müssen pro Punkt mindestens vier virtuelle Linienfeatures zu anderen Punkten übereinstimmen, die wiederum jeweils mindestens drei Kanten schneiden. Die Qualität der Zuordnungen jedoch bleibt gleich. Wie die Ergebnisse in Abbildung 6.15 zeigen, sind die gefundenen Zuordnungen nahezu alle richtig.

6.2.6 Fehlerhafte Eckpunkte

Im letzten Kapitel wurde der Einfluss von zu wenig Eckpunkten untersucht. Dabei hatten alle der vorhandenen Punkte der kleineren Teilmenge eine Entsprechung im anderen Bild. In diesem Kapitel wird untersucht, was geschieht, wenn ein Bild Eckpunkte enthält, für die es keine Entsprechung im anderen Bild gibt. Dies kann beispielsweise durch Veränderungen der Objekte selbst oder durch Bildrauschen entstehen. Um dies zu untersuchen, wurden Testbilder erstellt, die zusätzlich zu den realen Eckpunkten und Kanten weitere Eckpunkte enthalten, welche im zweiten Bild nicht vorhanden sind. In Abbildung 6.16 sind solche fehlerhaften Punkte und das Ergebnis der Zuordnung für diese beiden Bilder dargestellt. Hier wurden zusätzlich zu den 28

Eckpunkten, die das Bild im Original enthält 30 weitere Punkte innerhalb der Strukturen hinzugefügt.

In Abbildung 6.17 sind die Ergebnisse der Testreihe angegeben. Zu sehen sind acht Kurven für unterschiedlich geneigte Bilder. Auf der X-Achse ist jeweils der Anteil der Fehlerhaften Punkte aufgetragen die überzählig im Bild sind im Verhältnis zu den identischen Punkten im Bild, auf der Y-Achse der Anteil der richtig detektierten Punkte. Bei 28 Punkten mit einer Entsprechung im anderen Bild bedeutet 100 %, dass doppelt so viele (56) Eckpunkte im Bild zu sehen sind, von denen 28 kein Pendant im Anderen Bild haben. Entsprechend bedeutet 200 %, dass zusätzlich zu den 28 Punkten mit einem Pendant, 56 falsche Punkte im Bild sind, also insgesamt 84 Punkte.

Aus den Ergebnissen ist ersichtlich, dass bis zu 50% fehlerhafte Punkte keine nennenswerten Auswirkungen auf das Ergebnis haben. Alle Fehleranteile, die darüber liegen, reduzieren die richtigen Zuordnungen im Bild. Das bedeutet, dass erst bei einem deutlichen Ungleichgewicht zwischen identischen und nicht identischen Eckpunkten in beiden Bildern, keine erfolgreichen Zuordnungen mehr gefunden werden. Solange ein Großteil der Punkte übereinstimmt, werden fehlerhafte Punkte ignoriert und die Punkte können zuverlässig zugeordnet werden.

6.2.7 Doppelverhältnisse in den Linienfeatures

Die Doppelverhältnisse in den virtuellen Linienfeatures sind die projektiv invarianten Merkmale, die während der Laufzeit des Algorithmus am häufigsten abgefragt werden. Daher ist es hier besonders wichtig, dass dies schnell geschieht und dass redundante Informationen zuvor reduziert werden. Ein Grund, warum der neue entwickelte Algorithmus schneller als andere projektiv invariante Verfahren ist (siehe Kapitel 6.5.2), sind die reduzierte Anzahl an Kombinationen, die untersucht werden müssen. Zwei weitere Gründe sind: Zum einen werden die Vergleiche der Doppelverhältnisse nicht über deren numerische Größe durchgeführt, sondern die Größen werden als Zeichenkette in einer Liste gespeichert, und die Zeichenketten werden stattdessen miteinander verglichen, was deutlich schneller ist. Zum anderen wird die Anzahl der untersuchten Doppelverhältnisse zuvor reduziert (siehe Kapitel 4.2.2). Diese Reduktion erfolgt anhand der Nähe, die zwei Doppelverhältnisse numerisch zueinander besitzen. Die folgenden Untersuchungen zeigen die Einflüsse, die unterschiedliche Reduktionswerte sowohl bei der Rechenzeit, als auch bei der Anzahl und der Richtigkeit der Punktkorrespondenzen auf das Endergebnis haben. Mehrere Werte zwischen 0.05 und 0.4 wurden untersucht. Die Ergebnisse der Rechenzeit sind in Abbildung 6.18 dargestellt.

Sehr kleine Werte bedeuten, dass mehr Doppelverhältnisse in der Featurebeschreibung enthalten sind. Bei großen Werten sind deutlich weniger Merkmalsbeschreibungen in einem Linienmerkmal enthalten, falls deren numerische Größen ursprünglich

sehr ähnlich waren. Aus den Ergebnissen der unterschiedlichen Rechenzeiten ist ersichtlich, dass geringere Reduktionswerte in längeren Laufzeiten resultieren und dass zwischen den Reduktionswerten 0.1 und 0.4 keine Verbesserung mehr in der Gesamtlaufzeit erreicht wird.

In Abbildung 6.19 sind die Ergebnisse aufgetragen, welche die unterschiedliche Anzahl an Zuordnungen und deren Richtigkeit in Abhängigkeit zu den verwendeten Reduktionswerten angibt. In der durchgeführten Testreihe sind die gefundenen Punktkorrespondenzen für Reduktionswerte von über 0.1 zuverlässiger als Werte darunter. Die Werte der Laufzeit, sowie der Anzahl und Richtigkeit der Punkte zeigen, dass ein Reduktionswert von 0.1 eine gute Wahl darstellt.

6.2.8 Zusammenspiel von Nachbarschaftsgröße und P^2 -Invariante

Die für die Plausibilitätstests benutzten P^2 -Invarianten sind sensitiv auf die Größe des Bereichs, in denen die fünf Punkte verteilt sind. Die folgende Untersuchung testet unterschiedliche Bildgrößen mit einem jeweils bildfüllenden Objekt. Untersucht wird, wie sich diese auf die Anzahl und die Qualität der Ergebnisse auswirken. Benutzt werden die gleichen Testbilder wie in Kapitel 6.1, jedoch mit unterschiedlichen Auflösungen. Die Ergebnisse sind in den Abbildungen 6.20, 6.21, 6.22 und 6.23 dargestellt. Die Resultate zeigen, dass große Bilder mit großflächigen Übereinstimmungen besser gefunden werden können, als kleine Bilder mit entsprechend kleinen übereinstimmenden Bereichen. Bei unterschiedlich großen Bildern ist der limitierende Faktor jeweils das kleinere Bild.

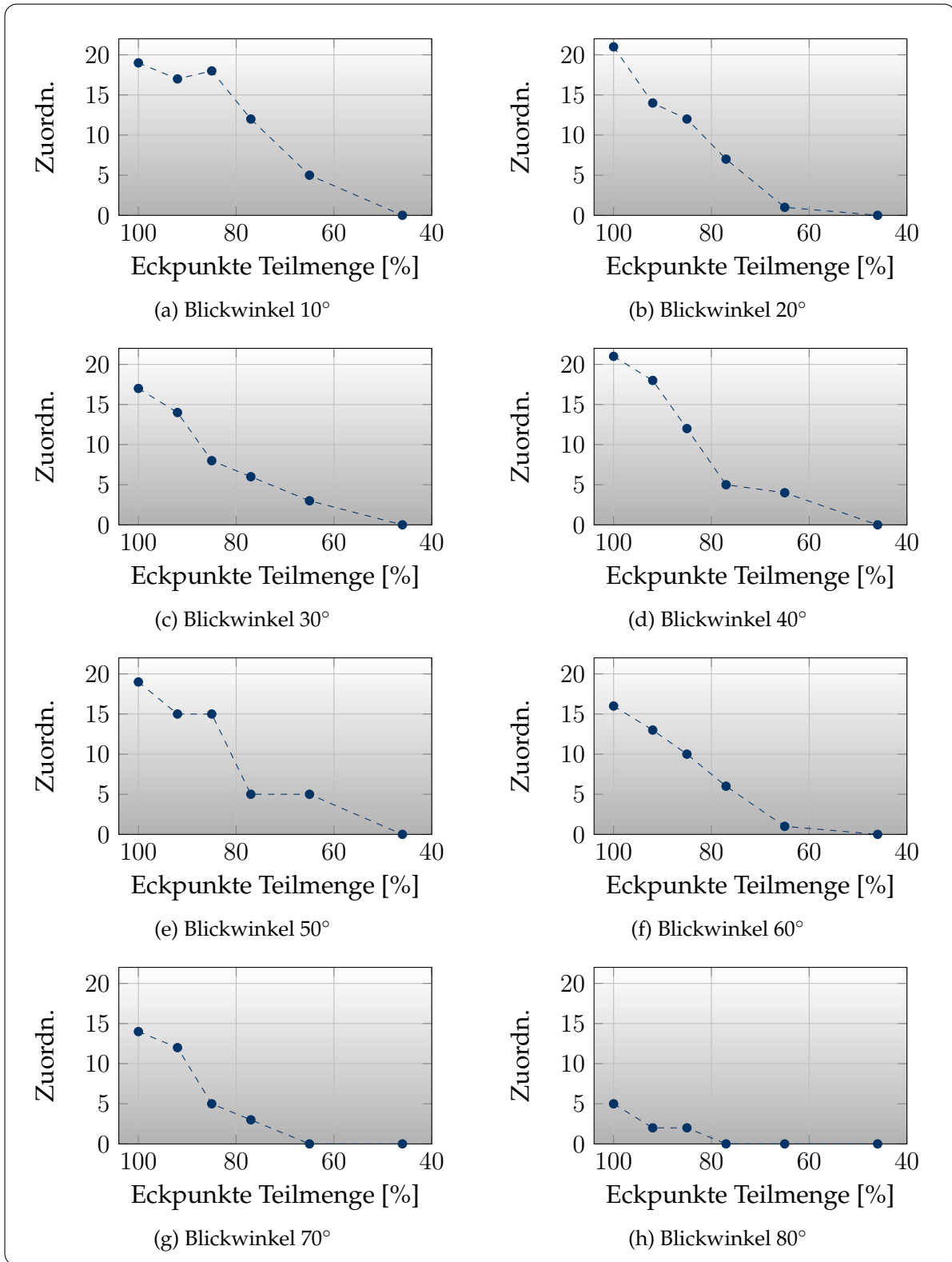


Abbildung 6.14: Anzahl Zuordnungen bei unterschiedlichen Teilbereichen.

6 Evaluation

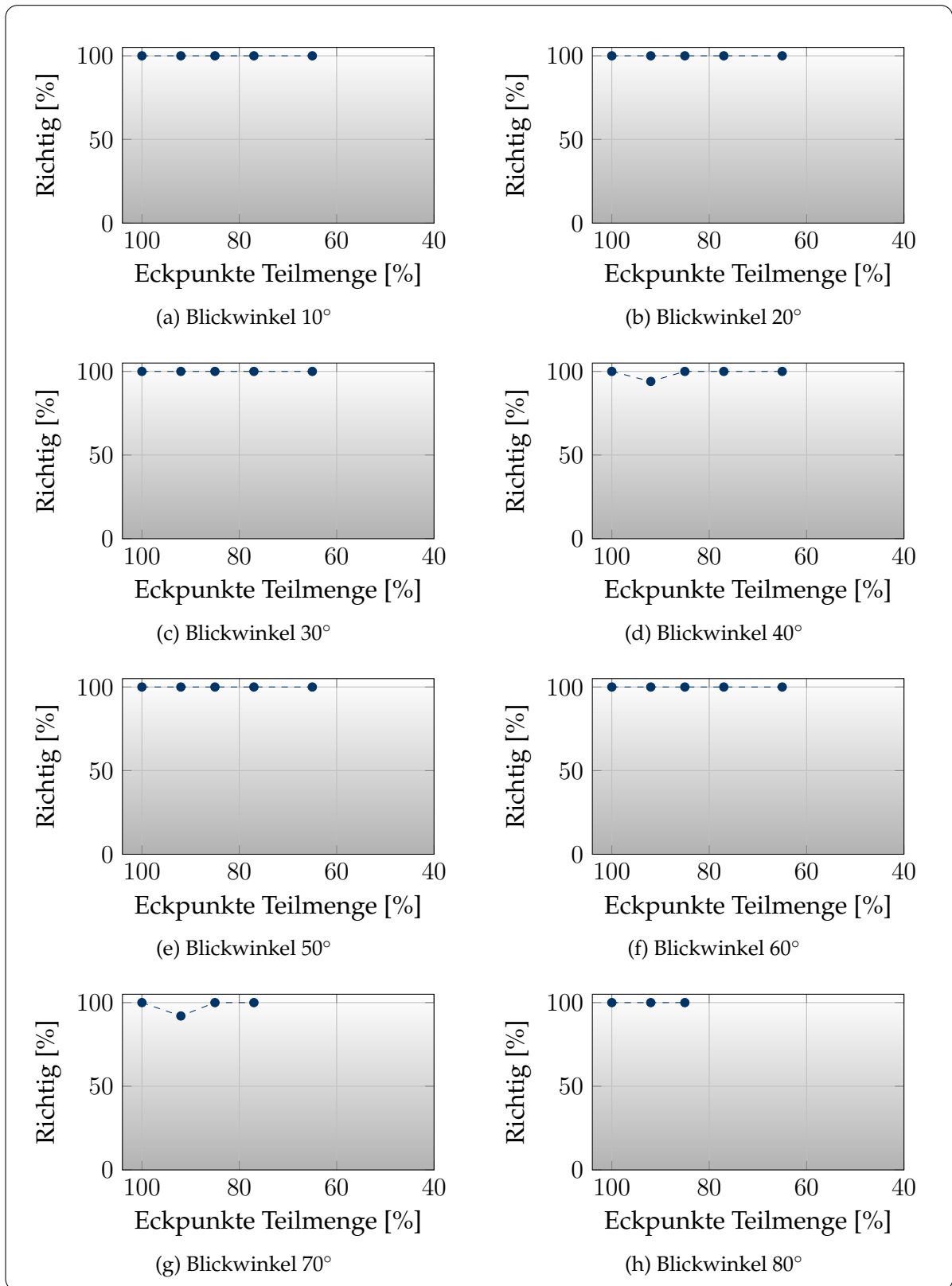


Abbildung 6.15: Richtige Zuordnungen bei Teilbereichen in Prozent.

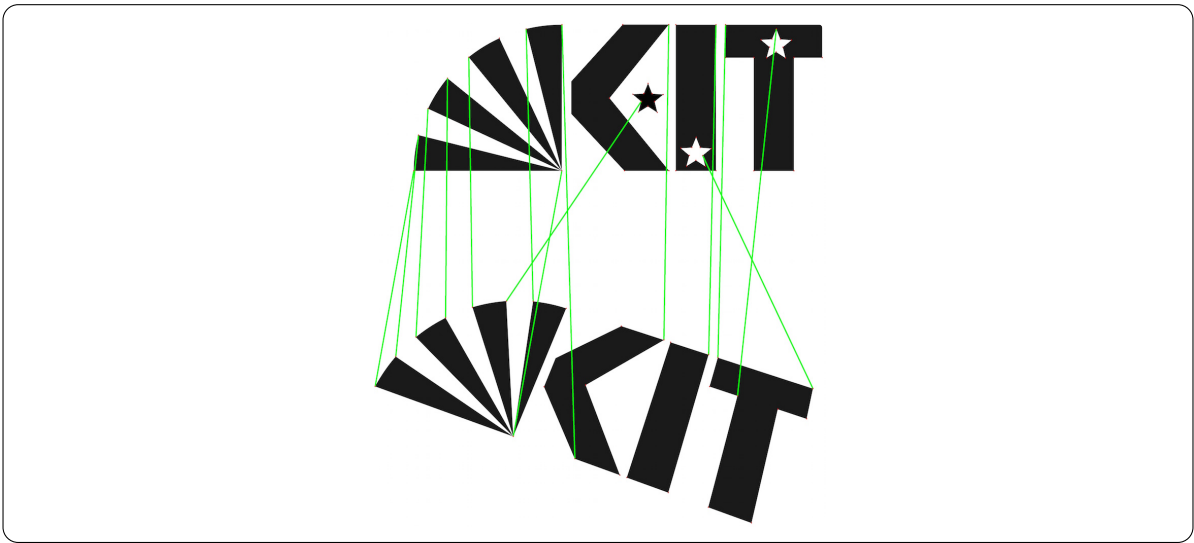


Abbildung 6.16: Unterschiedliche Eckpunkte in beiden Bildern. Diese können zu fehlerhaften Punktkorrespondenzen führen.

6 Evaluation

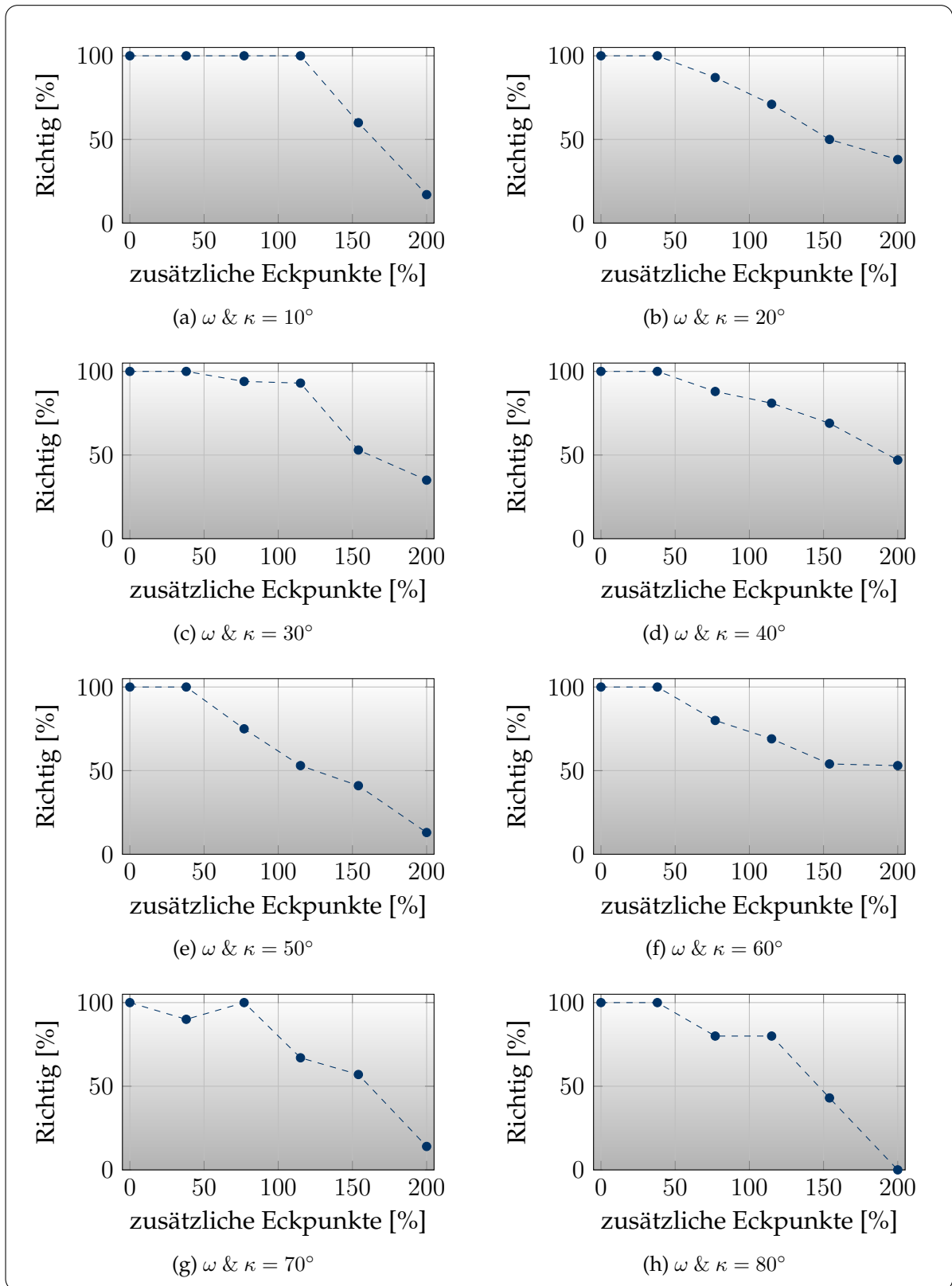


Abbildung 6.17: Richtige Zuordnungen in Abhängigkeit zu dem Anteil an fehlerhaften Punkten im Bild bei unterschiedlichen Blickwinkeln.

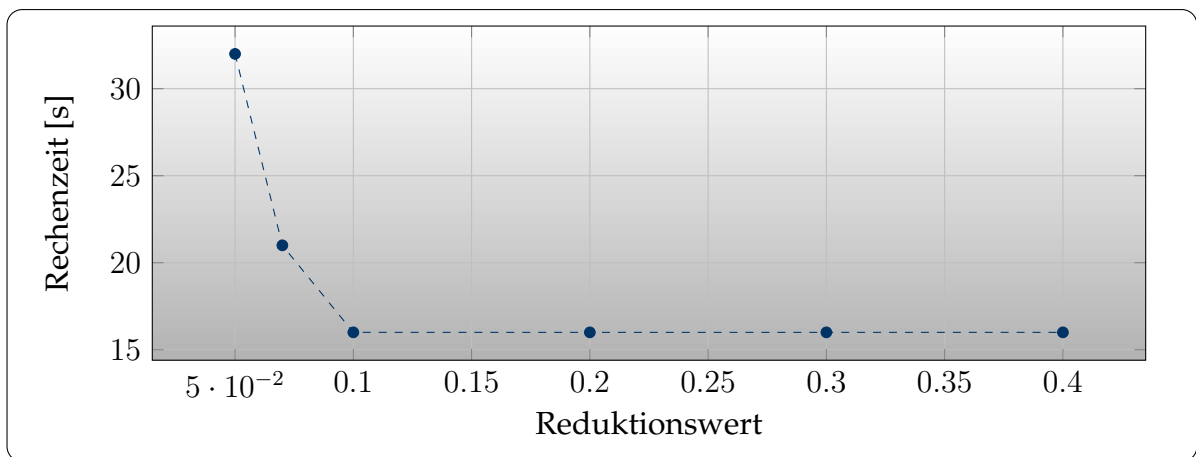


Abbildung 6.18: Einfluss des Reduktionswertes der Doppelverhältnisse auf die Rechenzeiten.

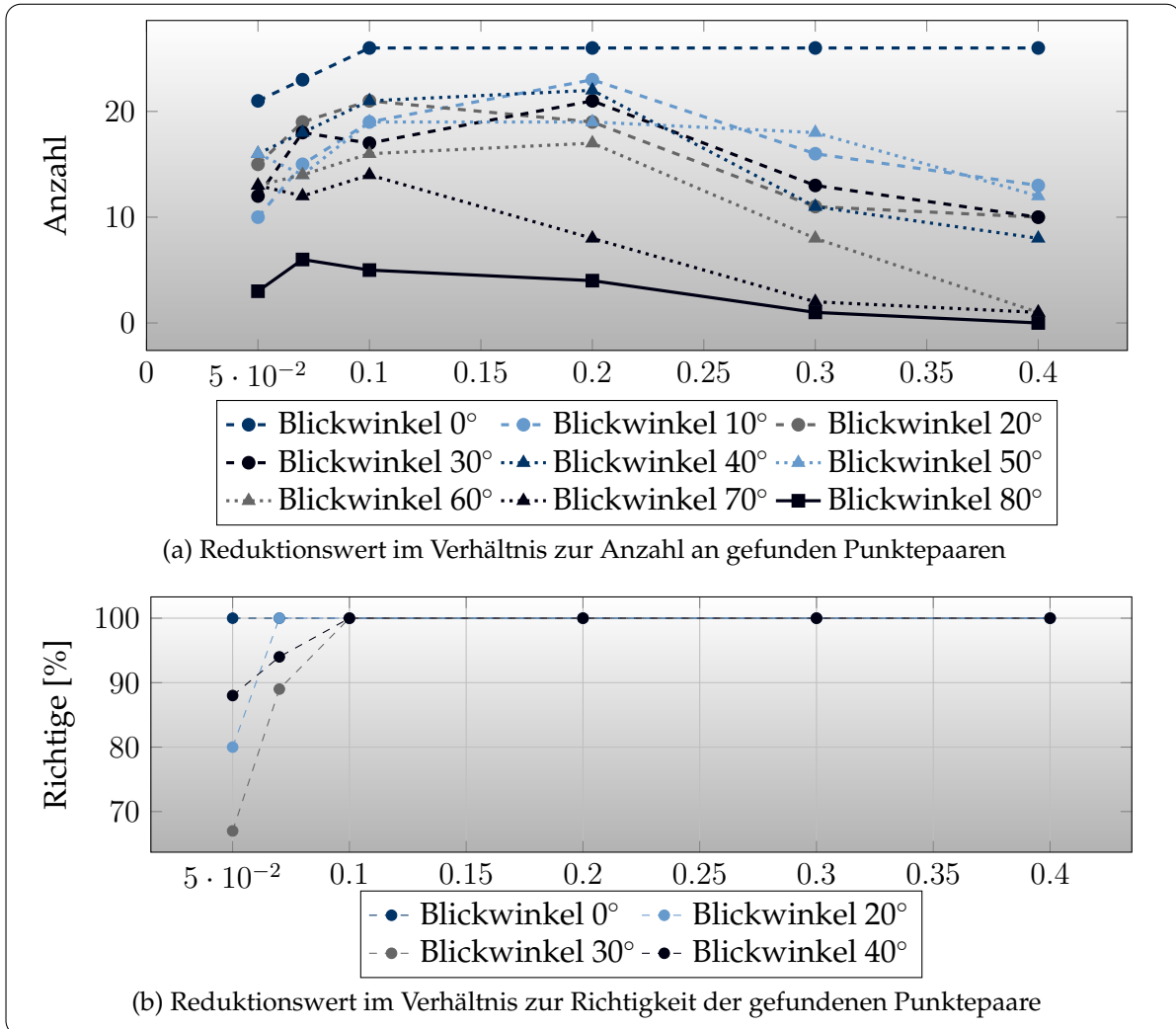


Abbildung 6.19: Einfluss des Reduktionswertes der Doppelverhältnisse auf die Anzahl und Richtigkeit der Punktzuordnungen.

6 Evaluation

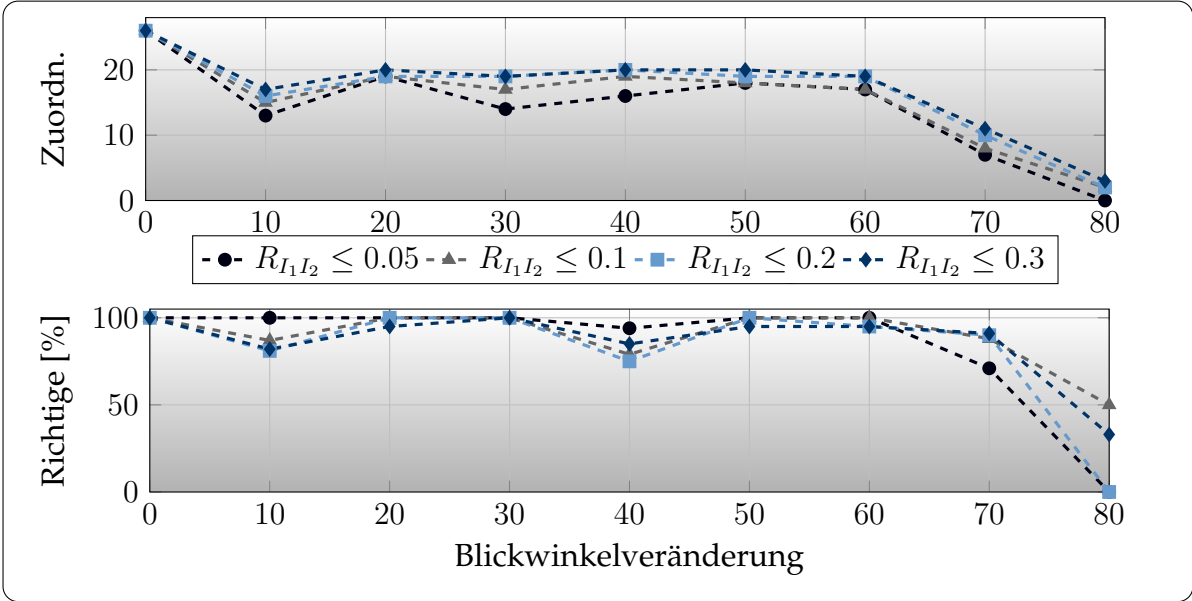


Abbildung 6.20: 2000 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.

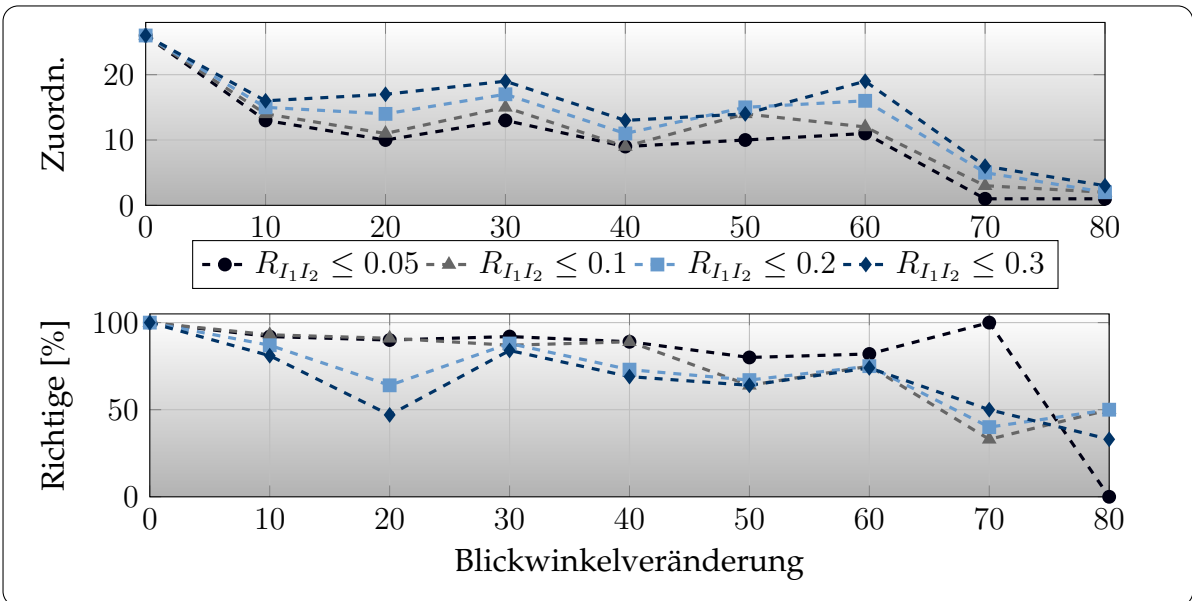


Abbildung 6.21: 1000 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.

6 Evaluation

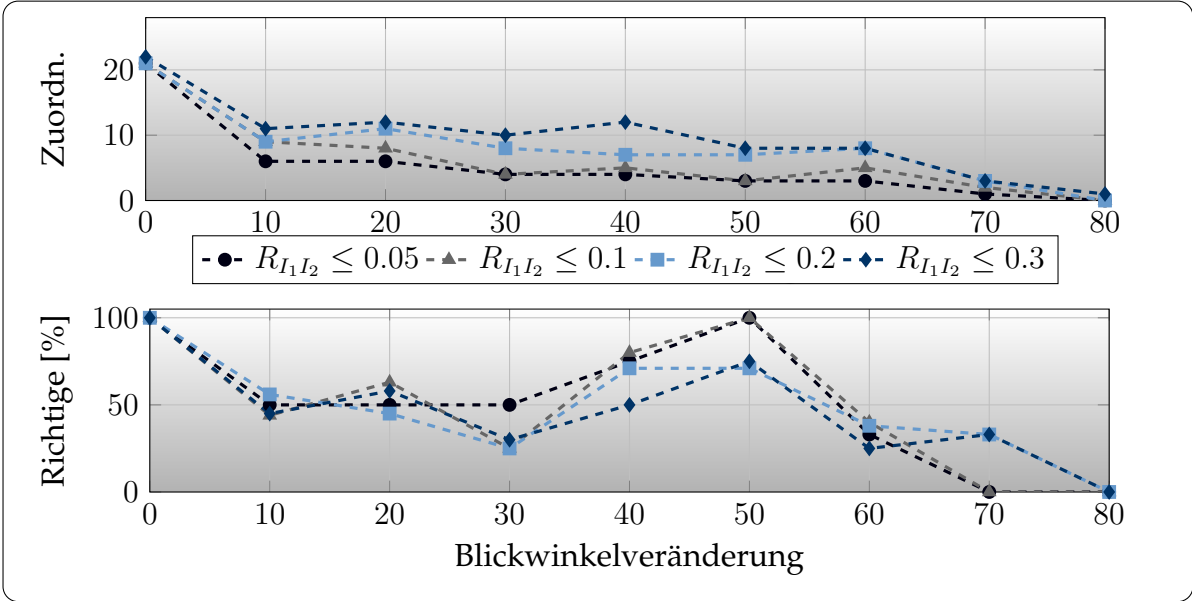


Abbildung 6.22: 500 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.

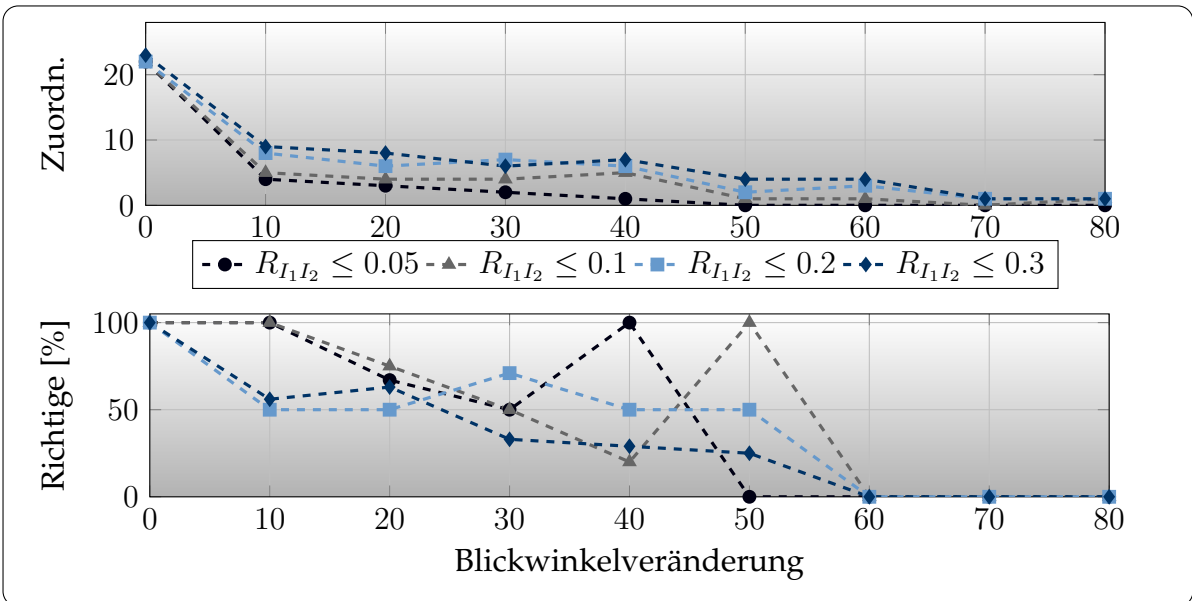


Abbildung 6.23: 250 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.

6.3 Übertragbarkeit der Verfahren und Realbilder

Ob ein Verfahren auf unterschiedliche Bereiche übertragbar ist, hängt davon ab, wie viele Annahmen getroffen wurden, die das Verfahren bei einigen Anwendungsbereichen glänzen lassen, aber bei anderen nicht praktikabel sind. Dieses Verfahren benutzt sehr wenig Annahmen, jedoch eine wesentliche: Durch die Verwendung der Doppelverhältnisse und der P^2 -Invariante I_1 und I_2 liegt der Anwendungsbereich bei ebenen Flächen.

Das Verfahren ist bewusst einfach gehalten und benutzt daher lediglich Eckpunkte und Kanten als Eingangsdaten. Dies hat beispielsweise folgende Vorteile: Im Gegensatz zu Kanten, die durch veränderte Schattenwürfe entstehen, bleiben Kanten, welche durch physikalische Tiefensprünge oder veränderte Materialeigenschaften entstehen, konstant. Diese sind weitgehend unabhängig von radiometrischen Veränderungen oder auch von unterschiedlichen Sensoren. Dadurch, dass nicht alle Kanten entlang der virtuellen Linien übereinstimmen müssen, sondern nur mindestens drei, können zum Beispiel variable Schattenkanten ignoriert werden. Durch diese Einfachheit lässt sich das neue Konzept auf andere Bereiche übertragen. Ein weiterer Vorteil liegt in der geringen Anzahl an Eingangsdaten, die zudem sehr einfach und schnell mit bestehenden Algorithmen zu berechnen sind. So können sowohl Indoor-Aufnahmen, als auch Outdoor-Aufnahmen verarbeitet werden. Bilder unterschiedlicher Kameras, die beispielsweise unterschiedliche Sensortypen oder unterschiedliche Farbtemperatureinstellungen haben, können verwendet werden. Nahezu beliebige Aufnahmerichtungen können abgedeckt werden (siehe Kapitel 6.2.2), die dazu noch unterschiedlich perspektivisch deformiert sein können (siehe Kapitel 6.2.3).

Um zu demonstrieren, dass das Verfahren auch bei Realdaten funktioniert, wurden folgende Aufnahmen als Testdatensätze verwendet (Abbildungen 6.24, 6.25 und 6.26). Die Zuordnungen wurden trotz der Verzeichnung der Bilder, die nicht korrigiert wurde, erfolgreich gefunden. Der Algorithmus wurde dafür jeweils auf einen Bildausschnitt angewendet. Die abgebildeten Visualisierungen zeigen das gesamte Bild, um einen besseren Eindruck der Szene und der unterschiedlichen Blickrichtungen zu bekommen.

6.4 Mehrere Ebenen

Um zu zeigen, dass das neu entwickelte Verfahren auch mit mehreren Ebenen im gleichen Bild funktioniert, wurde folgendes Bildpaar benutzt (siehe Abbildung 6.27).

Das Bild enthält zwei große Ebenen, nämlich die beiden weißen Blätter mit den Schriftzügen *EINS* und *ZWEI*. Zwischen beiden Aufnahmen wurde zum einen die Position der Kamera verändert und zum anderen wurden die beiden Ebenen relativ zueinander bewegt. Das Verfahren wurde jeweils auf das gesamte Bild angewandt. Die



Abbildung 6.24: Ticket-Automat aus zwei unterschiedlichen Blickwinkeln aufgenommen. Die perspektivischen Effekte sind an der Schrift gut zu sehen. Auch die farblichen Unterschiede in beiden Bildern sind gut sichtbar. 17 Zuordnungen.



Abbildung 6.25: Werbetafel eines Brezelstandes aus zwei unterschiedlichen Blickwinkeln fotografiert. 32 Zuordnungen, davon eine Fehlzusordnung.



(a) Werbetafel von L'TUR aus zwei unterschiedlichen Blickwinkeln fotografiert.



(b) Erfolgreiches Matching bei komplett unterschiedlichen radiometrischen Informationen. Einzig die zugrunde liegende Form des Logos ist dieselbe.

Abbildung 6.26: Weitere Realbeispiele.



Abbildung 6.27: Testbilder zum testen mehrerer Ebenen. Die Kameraposition und die Positionen der Ebenen wurden verändert

grünen Verbindungslinien in Abbildung 6.28 zeigen jeweils eine Punktzuordnung, welche wie beschrieben (siehe Kapitel 5) aus 5 Eckpunkten abgeleitet wurde. Jede Linie stellt daher eine eigenständige Transformation dar.

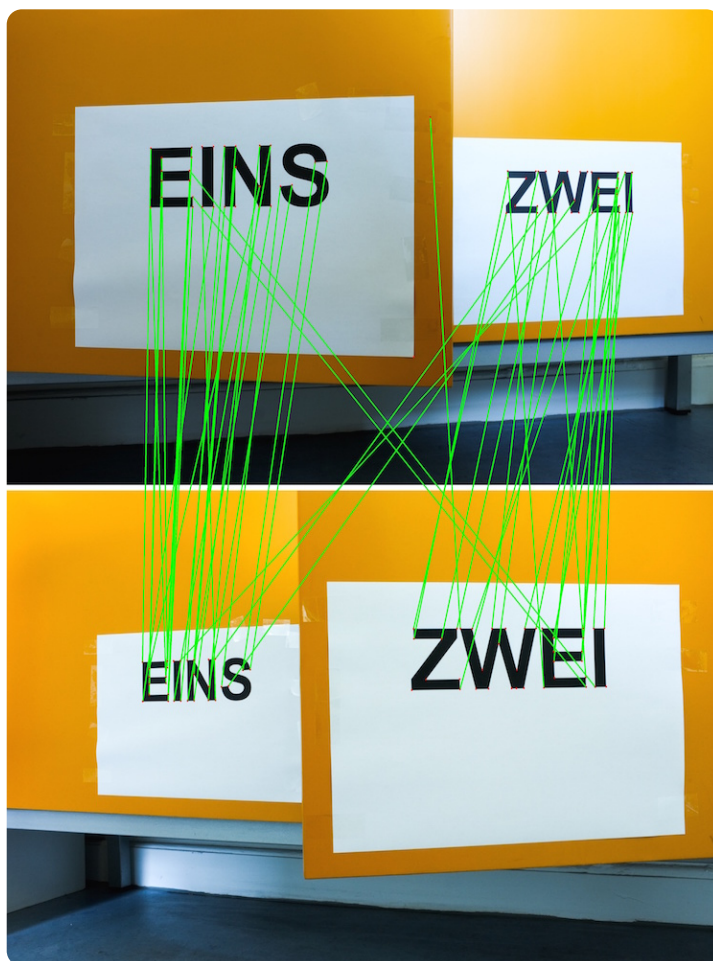


Abbildung 6.28: Zuordnungen bei unterschiedlichen Ebenen. Jede grüne Linie entspricht einer kompletten Transformation.

Zu sehen ist, dass das Verfahren auch bei unterschiedlichen Ebenen korrekte Ergebnisse liefert, auch wenn nicht alle Punktzuordnungen richtig sind. Eine weitere Visualisierung des Ergebnisses ist in Abbildung 6.29 zu sehen.

Hier wurde jeweils eine Transformation von einer Korrespondenz aus der ersten Ebene und eine Transformation von einer Korrespondenz aus der zweiten Ebene verwendet. Das erste Bild wurde mit einer Homographie so entzerrt, dass die beteiligten Punkte des ersten Bildes an der gleichen Position wie die Punkte des zweiten Bildes liegen. Das Resultat wurde wie ein Anaglyphenbild farblich kodiert. Die Abweichungen, die nach der Transformation sichtbar geworden sind, wurden in Blau und Rot dargestellt. Innerhalb der Ebene *EINS* sind die Abweichungen gering. In Ebene *ZWEI* sind die Differenzen etwas größer, was auf die Genauigkeit der Punkte-Extraktion in



Abbildung 6.29: Bildüberlagerung der aufeinander transformierten Ebenen. Links für identische Punkte aus Ebene *EINS*, rechts für Punkte aus Ebene *ZWEI*.

Kombination mit dicht beieinander liegenden identischen Punkten zurückzuführen ist. In beiden Fällen gelten die Transformationen selbstverständlich nur für die jeweilige Ebene. Die jeweils andere Ebene wird daher natürlich mit großen Abweichungen dargestellt.

6.5 Ergebnisvergleich mit anderen Verfahren

Um die Ergebnisse des in dieser Arbeit neu konzipierten Algorithmus mit denen anderer konkurrierenden Algorithmen zu vergleichen stellt sich zunächst die Frage, mit welchen es verglichen werden kann und soll.

6.5.1 Ergebnisvergleich: Blobbezogene Verfahren

Sehr häufig wird der SIFT-Algorithmus als Referenz bei Punktkorrespondenzen benutzt. Der SIFT-Algorithmus ist ohne Zweifel eines der bisher leistungsfähigsten Verfahren und vor allem eines der am häufigsten zitierten. Die Tatsache, dass SIFT so oft als Referenz benutzt wird, kann aber auch einem anderen Grund geschuldet sein, nämlich, dass der Autor bereits sehr früh eine funktionierende und sehr leicht zu benutzende ausführbare Version des Programms im Internet zur Verfügung gestellt hat. So konnte und kann jeder besonders leicht diesen Algorithmus testen und die Ergebnisse als Referenzergebnisse benutzen. Konzeptionell gesehen ist SIFT kein konkurrierendes Verfahren. Wie bereits bei der Veröffentlichung des Algorithmus beschrieben ([Lowe, 1999] (S.6) und [Lowe, 2004] (Figure 9 S.18)), fällt die Wiederholgenauigkeit bei Blickwinkelveränderungen über 20 Grad stark ab. Der direkte Vergleich mit SIFT ist daher sehr einfach. Alle zuvor gezeigten Realbild-Datensätze sind mit dem SIFT-Algorithmus nicht zuordenbar. Es werden keine richtigen Übereinstimmungen gefunden.

Der ASIFT-Algorithmus [Yu & Morel, 2009] ist von der Konzeption her kein projektiv invarianter Algorithmus, sondern wie der Name bereits sagt ein affin- und skaleninvarianter Algorithmus. Zudem ist er nicht robust gegenüber radiometrischen Veränderungen. Da mit seiner Hilfe aber auch sehr starke Blickwinkelveränderungen handhabbar geworden sind, kann er wegen dieser Fähigkeit als Referenz im Bezug auf unterschiedliche Blickwinkel gesehen werden. Mit ASIFT ist jedoch eine Zuordnung von Bildausschnitten mit starker perspektivischer Deformation nicht möglich, da diese mit dem benutzen affinen Kameramodell nicht abgebildet werden können, und zum anderen sind Zuordnungen von Bildern mit deutlich unterschiedlichen radiometrischen Informationen ebenfalls nicht möglich. So werden beispielsweise für das Bildpaar in Abbildung 6.26 (b) keine Punktkorrespondenzen gefunden.

6.5.2 Ergebnisvergleich: Projektiv invariante Verfahren

Die als projektiv invariant konzipierten Verfahren in der Literatur, welchen als Referenz dienen können, sind in Kapitel 2.2.2 vorgestellt. Diese sind [Meer *et al.*, 1994] und [Luo *et al.*, 2013]. Letztgenanntes ist ausgelegt für kleinräumige abgeschlossene Symbole und Buchstaben. Verdeckungen und Teilbilder sind hier nicht handhabbar. Der Anwendungsbereich ist daher zu speziell und nicht allgemein genug gehalten. Der Algorithmus von [Meer *et al.*, 1994] hat, wie in Kapitel 3.3.1 besprochen, den großen Nachteil, dass der Rechenaufwand bedingt durch die $\binom{n}{k}$ mit $k = 5$ Rechenoperationen, die je n Punkte notwendig sind, extrem hoch ist. Zudem werden bei dieser als *blinde Technik* bezeichneten Vorgehensweise lediglich Punktkoordinaten als Eingangsdaten genutzt, was beispielsweise eine abschließende automatische Bestätigung, wie im vorgestellten neuen Algorithmus ausschließt (siehe Kapitel 5.2.4). Ein deutlicher Unterschied zu dem reinen Vergleichen von P^2 -Invarianten ist zudem, dass mit dem Algorithmus dieser Arbeit Punkte zugeordnet werden, die mehrfach auf unterschiedliche Art und Weise bestätigt wurden. Im Gegensatz dazu werden bei [Meer *et al.*, 1994] alle Punkte zugeordnet, sobald der Test der P^2 -Invarianzen zumindest einmal je Punkt eine Bestätigung ausgibt. Das ist unabhängig davon, ob der Punkt tatsächlich eine Entsprechung im anderen Bild hat oder nicht. Dies führt bei bisherigen Algorithmen im Fall von Verdeckungen und einer daraus resultierenden unterschiedlichen Anzahl an Punkten zwischen beiden Bildern zu falschen Ergebnissen. Das in dieser Arbeit neu erarbeitete Konzept hat den Vorteil gegenüber anderen projektiv invarianten Algorithmen, dass es mit deutlich weniger Rechenoperationen auskommt und Verdeckungen und unterschiedliche Bildausschnitte handhabbar werden. Dieser Vorteil kommt durch die Kombination aller invarianten Kenngrößen in einem Algorithmus zustande. Im Gegensatz zu anderen Ansätzen, die jeweils nur auf einem einzigen invarianten Merkmal aufsetzen, wie etwa der P^2 -Invariante, können hier durch die Verbindung aller invarianten Merkmale zuverlässigere und schnellere Ergebnisse erzielt werden. An jedem Punkt werden für den Anfangstest nicht mehr wie bei älteren Verfahren $\binom{n}{k}$ mit $k = 5$ Rechenoperationen durchgeführt, sondern nur noch $\binom{n}{k}$

mit $k = 2$ Rechenoperationen. Bei einem Punktevergleich von $n_1 = 18$ Punkten im ersten Bild und $n_2 = 21$ Punkten im zweiten Bild entspricht das beispielsweise einer Reduzierung von 174 350 232 untersuchten Kombinationen im Fall von [Meer *et al.*, 1994] auf gerade mal 32 130 im vorliegenden Fall. Der Beschleunigungsfaktor durch die unterschiedlichen Anzahl an Kombinationen für den Anfangstest beträgt in diesem Beispiel 5 426,4. Da nach diesem Anfangstest noch die neuen Plausibilitätstests durchgeführt werden, verringert sich dieser Faktor. Mit dem neuen hier vorgestellten Algorithmus betrug die Berechnungszeit für eine Punktzuordnung zweier Bilder mit 18 und 21 Punkten ca. 7 Sekunden. Ein Durchlauf der zwei gleichen Bilder mit dem Algorithmus von [Meer *et al.*, 1994] dauerte 1 h 26 min 41 s, was somit das 743-fache der Zeit ist.

6.5.3 Konzeptionell bedingte Unterschiede

Ein weiterer Unterschied zu den Algorithmen, die auf Punkte oder Blobs bezogen sind, liegt darin, dass bedingt durch die Benutzung von Doppelverhältnissen entlang ebener Linien und P^2 -Invarianten auf Flächen nur ebene Merkmale zugeordnet werden können. Beim reinen Stereomatching mit nur zwei Bildern können daher keine beliebig im Raum verteilten Punkte für die Zuordnung gefunden werden, sondern stattdessen viele unterschiedliche ebene Flächen. Wenn mindestens drei Bilder des gleichen Objektes verwendet werden, sind jedoch weitere Lösungsmöglichkeiten denkbar (siehe Kapitel 8).

Auch beziehen sich alle projektiv invarianten Eigenschaften auf zentralperspektivische (gnomonische) Abbildungen, bei denen die Kollinearitätsgleichungen gelten. Bilder, die mit Fischaugenobjektiven aufgenommen wurden, haben davon unterschiedliche Abbildungsvorschriften, wie etwa eine winkeltreue, eine äquidistante, eine flächentreue oder eine orthografische Projektion (vergleiche [Schneider *et al.*, 2009]). Diese unterschiedlichen Abbildungsvorschriften haben jeweils andere invariante Eigenschaften, die sich von denen projektiver zentralperspektivischer Abbildungen unterscheiden. Daher kann das hier neu erarbeitete Konzept nicht direkt bei Bildern aus Fischaugenkameras verwendet werden. Die Bilder müssten dafür erst in eine zentralperspektivische Abbildung transformiert werden.

Ein weiterer Unterschied liegt in der Art des Merkmalvergleichs zwischen beiden Bildern. Im Gegensatz zu den gängigen Algorithmen, die bezogen sind auf Einzelpunkte oder Blobs, werden bei dem Verfahren nicht alle Merkmale im Vorfeld berechnet und danach miteinander verglichen. Lediglich die Linienmerkmale für die Anfangshypothesen werden im Voraus erstellt. Die anderen vergleichbaren Merkmale, wie die Nachbarschaftsreihenfolgen, die konvexen Hüllen von Punkten und die P^2 -Invarianten werden erst zur Laufzeit berechnet, wenn aufgrund der Hypothesen durch die Linienmerkmale an einem Punkt dazu Veranlassung besteht, diese

6 Evaluation

zu vergleichen. Dies spart sehr viele Rechenschritte, verglichen mit anderen Verfahren, die alle ihre Merkmale im Vorfeld erstellen. Das Verfahren ist daher eines, das während der Laufzeit aktiv nach identischen Punkten sucht.

7 Zusammenfassung

Das vorgestellte neue Konzept dieser Arbeit bietet die Möglichkeit, ein sehr herausforderndes Problem in der *Photogrammetrie* und der *Computer Vision* zu lösen.

Ein wichtiges Ergebnis besteht darin, dass das Konzept der Punktzuordnung unter ausschließlicher Benutzung projektiver und radiometrischer Invarianzen funktioniert und auf Realdaten angewendet werden kann. Der Schlüssel zur Lösung liegt in der Kombination aus mehreren unterschiedlichen zur Verfügung stehenden Invarianzen. So werden alle als projektiv invariant herausgestellten Merkmale (diese sind: gerade Linien, Inzidenzen, Doppelverhältnisse, P^2 -Invarianzen und Konnektivitäten) in Kombination bei der Suche nach Punktzuordnungen benutzt. Die Ergebnisse zeigen, dass der Algorithmus korrekte Punktkorrespondenzen sogar bei starken perspektivischen Veränderungen findet, was Unterschiede des Blickwinkels und veränderte Kamerakonstanten mit einschließt. Durch den allgemein gehaltenen Lösungsansatz ist die Übertragbarkeit des neuen Konzeptes sehr hoch. Es stellt keine Insellösung dar. Weitere Alleinstellungsmerkmale ergeben sich daraus, dass der Algorithmus auch bei partiellen Verdeckungen oder Bildausschnitten sinnvolle Ergebnisse liefert, einfache Rechenoperationen benutzt und anders als bei den vorhandenen Lösungsansätzen deutlich weniger Rechenoperationen benötigt als diese. Im Gegensatz zu anderen denkbaren Lösungsansätzen wird kein Modellwissen über die Szene benutzt. Es müssen lediglich virtuelle Linien und Ebenen vorhanden sein. Alle Informationen werden direkt aus den Bildern abgeleitet. Der dreidimensionale Raum wird zu keinem Zeitpunkt explizit genutzt, was bedeutet, dass im gesamten Verfahren keine dreidimensionalen Objektkoordinaten benutzt werden.

Die Neuheit und das wesentliche Alleinstellungsmerkmal der Arbeit bestehen in der Beschreibung eines Bildes mittels mehrerer kombinierter projektiv invarianter Eigenschaften. Die Verwendung einzelner projektiv invarianter Eigenschaften zum Bestätigen identischer Bildpunkte ist nicht neu. Jedoch wurden bisher nicht die verschieden projektiv invarianten Eigenschaften kombiniert. Diese Kombination führt zu völlig neuen Möglichkeiten der Bildkorrespondenzsuche. Ein weiterer Vorteil dieses Lösungsansatzes bedeutet die Anwendbarkeit auch auf Bilder, die mit einfacheren Transformationen abgebildet wurden, da die projektiv invarianten Eigenschaften auch für diese Transformationen gelten wie beispielsweise eine Translation, eine Rotation, eine Skalierung oder eine affine Abbildung.

Zusammengefasst sind die Neuheiten und Alleinstellungsmerkmale, die das Konzept des in dieser Arbeit vorgestellten Algorithmus für die Suche nach Punktkorre-

7 Zusammenfassung

spondenzen auszeichnen, die Kombination aus folgenden erreichten Zielen:

1. Das neue Konzept ermöglicht projektiv invariantes Matching bei zentralperspektivischen Abbildungen. Dazu zählen auch extrem starke Blickwinkelunterschiede zwischen beiden Bildern.
2. Das neue Konzept ermöglicht es, Bildpunkte selbst dann zuzuordnen, wenn zusätzlich zur Veränderung des Blickwinkels eine unterschiedliche Kamerakonstante verwendet wurde, was sich in unterschiedlichen optischen Verkürzungen im Bild auswirkt. Vorwissen über die verwendete Kamerakonstante ist dabei nicht notwendig.
3. Das neue Konzept ermöglicht es, Bilder mit starken Beleuchtungsunterschieden oder unterschiedlichen Sensortypen zuzuordnen.
4. Das neue Konzept ist übertragbar und allgemein anwendbar in unterschiedlichsten Anwendungsbereichen. Das Prinzip ist dafür bewusst allgemein gehalten.
5. Das neue Konzept kann mit Teilausschnitten und Verdeckungen umgehen. Bisherige projektiv invariante Algorithmen bieten lediglich 1:1 Zuordnungen mit der gleichen Anzahl an Punkten im ersten wie im zweiten Bild.
6. Im Vergleich zu anderen Lösungsansätzen, welche projektiv invariante Einzelpunktzuordnungen anbieten, ermöglicht die Einführung der neu konzipierten Zuordnung von Linienmerkmalen eine drastische Reduzierung von Rechenschritten.

8 Ausblick und Vorschläge für zukünftige Arbeiten

Jede Arbeit, die neue wissenschaftliche Erkenntnisse hervorbringt, fördert automatisch auch neue Ideen, neue Ansatzmöglichkeiten und neue theoretische Lösungswege zutage, welche bis dahin entweder nicht bekannt waren oder vormals noch keine Relevanz hatten. In diesem abschließenden Kapitel werden Vorschläge für zukünftige Arbeiten zusammengefasst, die aus aufgetretenen Fragestellungen während der Arbeit entstanden sind. Als Ziel wird dabei ein Konzept gesehen, das in jedem Bildpaar identische Punkte findet, in dem auch ein Mensch identische Punkte erkennt. Die folgenden Ideen bauen auf den Ergebnissen der vorliegenden Arbeit auf und können für die Erstellung zukünftiger Algorithmen in Folgearbeiten hilfreich sein.

Das hier vorgestellte neue Konzept ist bewusst allgemein gehalten, um zu zeigen, dass es auf viele Bereiche übertragbar ist. Um den Zuordnungsprozess weiter zu beschleunigen, können zusätzliche Bildinformationen mit berücksichtigt werden, die zwar nicht radiometrisch invariant sind, aber bei gleichbleibenden äußeren Bedingungen sehr gute und vor allem sehr schnelle Ergebnisse liefern. So können zusätzlich zu den Doppelverhältnissen auch die Richtungen der Gradientenamplitudensprünge an den Kanten als Featurebeschreibungen benutzt werden. Dadurch bekommt die Linie eine eindeutige Richtung, sowie deutlich mehr Information, anhand derer sie zuordenbar ist. Allerdings wird dadurch die radiometrische Invarianz eingeschränkt. Erste Versuche diesbezüglich waren vielversprechend.

Um die Plausibilitätstests weiter zu beschleunigen, können zusätzliche einfache Schwellwerte eingeführt werden, welche die benutzten Nachbarn auf diejenigen mit den X meisten Übereinstimmungen bei den Linienmerkmalen reduzieren. Dies ermöglicht eine deutlich schnellere Überprüfung für den Fall, dass sehr viele Linienübereinstimmungen an einem Punkt gefunden wurden.

Ein anderer Ansatz wäre, einen projektiv invarianten Vermaschungsalgorithmus zu entwerfen, mit dem die X nächsten Nachbarn gefunden werden können. Das würde es ersparen, alle Punktkombinationen zu testen. [Chotikakamthorn *et al.*, 2005] bietet dafür einen guten Ansatz. Der Algorithmus kann projektiv invariante Vermaschungen durchführen, allerdings ist er abhängig von den benutzten Feature-Werten der Feature-Detektoren und ist daher nicht radiometrisch invariant.

Eine ausgezeichnete Möglichkeit, eine initiale Nachbarschaft zu finden, die in beiden Bildern gleich ist, kann bei [Criminisi, 2001] gefunden werden. Hier werden ebene

Bereiche im Bild selektiert, was auch als Beschränkung bei der Suche nach Nachbarpunkten benutzt werden kann.

Im gesamten Verfahren wird bewusst kein statistisches Verfahren, wie zum Beispiel RANSAC verwendet, um zu zeigen, dass erfolgreiche Punktzuordnungen auch aufgrund von logischen Beziehungen gefunden werden können (vergleiche Kapitel 5.2.4). Selbstverständlich kann in das Verfahren an mehreren Stellen ein solcher Test eingebaut werden, der überprüft, ob die gefundenen Punkte mit einer Homographie vom einen ins anderen Bild übergeführt werden können. Dies kann die Zuverlässigkeit und auch die Schnelligkeit nochmals erhöhen.

Die Schnittpunkte der Linienfeatures mit dem Kantenbild werden aktuell nur zur Bestätigung der Doppelverhältnisse benutzt. Falls das Linienfeature jedoch zwischen zwei Punkten liegt, die als Korrespondenzen bestätigt wurden, so können auch die Schnittpunkte des Linienfeatures als zusätzliche identische Merkmale verwendet werden. Diese Merkmale können beispielsweise beim dichten Matching von Bildern mit sehr starken Blickwinkelunterschieden helfen, identischen Punkte an texturarmen Flächen zu finden.

Wie beschrieben ist die Kombination aus mehreren Invarianzen der Schlüssel für eine erfolgreiche Punktzuordnung. Zu den Invarianzen gehört das Doppelverhältnis aus vier Punkten auf einer Geraden. Daneben gibt es noch eine Reihe weiterer Doppelverhältnisse im eigentlichen Sinn des Wortes. Ein Doppelverhältnis ist das Verhältnis von Verhältnissen. Dies lässt sich auf andere Kenngrößen im Bild übertragen, wie zum Beispiel Winkel oder die invarianten Eigenschaften die bei [Coelho *et al.*, 1992] beschrieben werden. [Quan, 1995] beschreibt, wie sechs Punkte im Raum als Invarianzen benutzt werden können, wenn mindestens drei Bilder des gleichen Objektes verwendet werden. Der Vorteil davon ist, dass diese sechs Punkte nicht länger auf ebene Flächen beschränkt sind wie die anderen invarianten Eigenschaften. Viele unterschiedliche Kombinationen dieser invarianten Kenngrößen sind denkbar, die alle bisher nicht untersucht wurden. Es gibt einen weiteren Grund, warum es sich lohnt, auch in Zukunft einen Fokus auf projektive Abbildungen und deren unterschiedliche Invarianzen zu legen: Eine allgemein gültige Lösung des Zuordnungsproblems für projektive Transformationen löst automatisch auch das Zuordnungsproblem für die anderen untergeordneten Transformationen (vergleiche Tabelle 2.1 auf Seite 17). In diesem Sinne: *Schlage die Trommel und fürchte dich nicht!* [Heine, 1844].

Anhang

Abbildungsverzeichnis

1	Erfolgreich detektierte Punktkorrespondenzen aus projektiven und radiometrischen Invarianzen in einem Stereobildpaar mit extrem weiter Basis.	v
1.1	Ein Beispiel für einen stark veränderten Blickwinkel. Die starke perspektivische Transformation zwischen Teilbereichen der Bilder erschwert eine automatische Suche nach identischen Punkten in beiden Bildern. .	2
1.2	Ein Beispiel für veränderte Radiometrie: Die zwei Luftbildaufnahmen bilden das gleiche Gebiet in der Stadt Leicester in England ab. Durch unterschiedliche Lichtquellen und Schattenwürfe ist die Textur in beiden Bildern stark unterschiedlich, was eine automatische Suche nach Punktkorrespondenzen erschwert. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)	3
2.1	Grundriss einer Aufnahmesituation eines einfachen Objektes. Bei einer automatischen Auswertung mit aktuellen Techniken müssen außer den vier Aufnahmerichtungen (durchgezogene Linien), welche notwendig wären, um jeden Punkt des Objektes in mindestens zwei Bildern abzudecken, zusätzlich viele weitere Zwischenschritte (gestrichelte Linien) aufgenommen und ausgewertet werden, weil für aktuelle Algorithmen der Winkel zwischen den vier Aufnahmen zu groß ist, da die perspektivische Transformation des Teilbildes nicht berücksichtigt wird.	9
2.2	Unterschiedliche Blickwinkel verändern die Textur in beiden Bildern, siehe (c) und (d). Obwohl dasselbe Objekt abgebildet wird, können Algorithmen, welche identische Punkte über kleinräumige Features suchen, keine richtigen Ergebnisse liefern.	10

Abbildungsverzeichnis

2.3	Die Bilder der Kirche in Rio de Contas in Brasilien (auszugsweise links dargestellt) wurden mit einem sehr hohen Überlappungsbereich und mit jeweils minimaler Änderung des Blickwinkels gegenüber dem Objekt aufgenommen. Für eine erfolgreiche automatische relative Orientierung der Bilder zueinander waren sehr viele Bilder von zahlreichen Standpunkten notwendig. Die errechneten Orientierungen der Bilder sind in den Bildern auf der rechten Seite in rot und grün visualisiert. Die Menge an Standpunkten ist notwendig, um projektive Transformationen im Bild möglichst klein zu halten und eine korrekte Punktzuordnung mit Hilfe des SIFT-Algorithmus durchführen zu können.	11
2.4	Ein Bildbeispiel für einen stark veränderten Blickwinkel zwischen zwei Kameraaufnahmen desselben Objektes. Mit lokalen Merkmalsdeskriptoren wie beispielsweise SIFT können keine korrekten Punktzuordnungen gefunden werden, da sich die kleinräumigen Kenngrößen nach einer projektiven Transformation stark unterscheiden.	12
2.5	Eine schematische Darstellung der Transformationen der jeweiligen Geometrien. A: Euklidische Transformationen, B: Ähnlichkeitstransformation, C: Affine Transformation, D: Projektive Transformation.	13
2.6	Die euklidische Transformation bestehend aus Translation und Rotation. Die invarianten Eigenschaften der euklidischen Transformation sind: Entfernungen, Winkel, Verhältnisse von Entfernungen, Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.	14
2.7	Die Ähnlichkeitstransformation bestehend aus Translation, Rotation und einer isotropen Skalierung. Die invarianten Eigenschaften der Ähnlichkeitstransformation sind: Winkel, Verhältnisse von Entfernungen, Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.	15
2.8	Die affine Transformation bestehend aus Translation, Rotation, Skalierungen entlang der beiden Achsen und einer Scherung. Die invarianten Eigenschaften der affinen Transformation sind: Parallelitäten, geometrische Schwerpunkte, Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.	15
2.9	Die projektive Transformation, bestehend aus Translation, Rotation, Skalierungen entlang der beiden Achsen, einer Scherung und der perspektivischen Projektion. Die invarianten Eigenschaften der affinen Transformation sind: Inzidenzen, Doppelverhältnisse, gerade Linien und Konnektivitäten.	16
2.10	Beispiele für unterschiedliche radiometrische Bildinformation bei verändertem Aufnahmezeitpunkt. Es ändern sich unter anderem: die Intensitäten der Farbkanäle, Schattenverläufe und Helligkeiten. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)	22

Abbildungsverzeichnis

3.1	Die gerade Linien der Kanten eines Lineals bleiben gerade Linien bei einer perspektivischen Projektion. Dies ist unabhängig von der Blickrichtung auf das Objekt.	25
3.2	Geometrische Inzidenz: Auftreten des Punktes P_3 auf der Strecke $\overline{P_1P_2}$. Wenn ein Punkt innerhalb der Strecke zwischen zwei räumlichen Koordinaten liegt, so liegt dieser Punkt auch nach einer perspektivischen Transformation noch innerhalb derselben transformierten Strecke. . . .	26
3.3	Geometrische Inzidenz: Auftreten mehrerer Punkte P_1, P_2, P_3, P_4 und P_5 innerhalb der Ebene E . Wenn ein Punkt in Ebene E auftritt, so tritt dieser Punkt auch nach einer perspektivischen Transformation $(P'_1, P'_2, P'_3, P'_4, P'_5)$ noch auf der transformierten Ebene E' auf. Liegt ein Punkt innerhalb einer konvexen Hülle, so liegt der Punkt auch nach einer perspektivischen Projektion innerhalb dieser transformierten Hülle.	26
3.4	Bei einer projektiven Transformationen ändern sich die Abstände zwischen den einzelnen Punkten auf der Geraden.	28
3.5	Die Invariante I_1 und I_2 von fünf identischen Punkten auf einer projektiven Ebene, sind im Originalbild und im transformierten Bild gleich.	29
3.6	Die Konnektivität von Punkten im Bild bleibt bei projektiven Transformationen erhalten. Die Reihenfolge der Nachbarn von einem Punkt aus gesehen bleibt gleich und die Reihenfolge der Punkte auf der konvexen Hülle bleibt gleich.	30
3.7	Starke Unterschiede in der radiometrischen Bildinformation des gleichen Objektes. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)	31
3.8	Gegenüberstellung der ursprünglichen RGB-Bilder und der Kantenbilder aus zwei unterschiedlichen Algorithmen. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)	33
3.9	Beispiele für radiometrisch invariante Eckpunkt-Features. Nicht alle Eckpunkte stimmen im Tag- und im Nachluftbild überein. Entscheidend ist jedoch, dass einige der Punkte übereinstimmen, die an markanten Stellen in beiden Bildern detektiert werden. Dies trifft zum Beispiel auf die Ecken an den Enden der großen mandelförmigen Glasfenster zu. (Copyright der Luftbilder: Mit freundlicher Genehmigung von Bluesky International Limited.)	34
3.10	Eine Verdeutlichung, warum Vermaschungen auf Grundlage von nächsten Nachbarschaften und Schwerpunkten nicht perspektivisch invariant sind. Ein Algorithmus, der die nächsten sechs Punkte abfragt, gibt im linken Bild die grünen Punkte zurück, im rechten, um 80 Grad geneigten Bild, die roten Punkte.	37
4.1	Als Markante Punkte im Bild werden Eckpunkte benutzt. Hier sind exemplarisch einige der Eckpunkte mit roten Punkten im Bild markiert.	39

Abbildungsverzeichnis

4.2	Zwischen den Eckpunkten werden die virtuellen Linienmerkmale erstellt. Die Schnittpunkte der Linie mit dem Kantenbild B und C bilden zusammen mit den Anfangs- und Endpunkten der Linie A und D vier Punkte auf einer geraden Linie. Von diesen vier Punkten kann nun das Doppelverhältnis ausgerechnet werden. Wenn die Punkte in beiden Bildern dieselben sind, so ist auch dieses Doppelverhältnis gleich.	40
4.3	Beispiel von mehreren Doppelverhältnissen entlang der virtuellen Linien. Die Doppelverhältnisse entstehen durch den Linienanfangs- und Endpunkt, sowie durch die dazwischen liegenden Schnittpunkte mit dem Kantenbild.	41
4.4	Der Aufbau der neu konzipierten projektiv und radiometrisch invarianten Linienmerkmale.	43
4.5	Berechnung des Richtungswinkels zwischen untersuchtem Punkt und Endpunkt des Linienfeatures.	44
4.6	Konzept zum Auffinden von Punktkorrespondenzen mit den neuen virtuellen Linienfeatures.	45
5.1	Voraussetzungen für die Plausibilitätstests: Es müssen vier Nachbarpunkte über die virtuellen Linienmerkmale bestimmt worden sein wie in Kapitel 4.3	47
5.2	Plausibilitätstests bezüglich konvexer Hülle und P^2 -Invarianten mit den gleichen fünf Punkten durchführen. Die Anzahl der Punkte auf der Konvexen Hülle, die Reihenfolge der Punkte, sowie die Invariante I_1 und I_2 müssen übereinstimmen, wenn die fünf Punkte die gleichen sind.	48
5.3	Abschließende Überprüfung: Punktkorrespondenz A (grün dargestellt) ist Teil der Punktkorrespondenzen, die nach den Plausibilitätstests die höchste Anzahl an Bestätigungen haben. Von diesen wird vermutet, dass darin zumindest eine Korrespondenz enthalten ist, welche richtig ist. Punktkorrespondenz B (blau gestrichelt) wird überprüft, indem die Linienfeatures in beiden Bildern, die von B nach A gehen (gelb dargestellt), übereinstimmen. Ist dies wie im Bild der Fall, wird die Punktkorrespondenz als richtig bewertet. Punktkorrespondenz C ist ein Beispiel für eine falsche Korrespondenz (rot gepunktet). Die Linienfeatures von C nach A (violett dargestellt) sind nicht identisch, weshalb die Punktkorrespondenz keine Bestätigung erhält. Nacheinander werden alle anderen Kandidaten, die bei den Plausibilitätstests ein positives Ergebnis hatten, gegen alle als richtig eingestuften Korrespondenzen (z.B. Punktkorrespondenz A) geprüft.	50
5.4	Das neue Konzept der Kombination aus Linien- und Flächenmerkmalen.	51
6.1	Faktor P für die perspektivische Deformation.	53
6.2	Synthetisch produzierte perspektivische Projektionen des Ursprungsbildes (a) unter Benutzung der Drehwinkel ω und κ . Auszugsweise ist jeder zehnte Schritt dargestellt.	53

Abbildungsverzeichnis

6.3	Unterschiedliche Anzahl und Position an Eckpunkten je nach Detektor.	54
6.4	Anzahl gefundener Eckpunkte in Abhängigkeit vom Neigungswinkel.	55
6.5	Ergebnisse der Punktzuordnung 60 und 85 Grad. Hier wurden Harris-Eckpunkte verwendet. Jede grüne Linie entspricht dabei einer vollständigen bzw. eindeutigen projektiven Transformation, da die Korrespondenz auf mindestens fünf Punkten beruht. Hervorzuheben ist, dass für dieses Ergebnis, wie auch für alle anderen, keine statistischen Ausreißertests notwendig waren.	56
6.6	Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.	57
6.7	Richtige Zuordnungen bei unterschiedlichen zugelassenen Spannweiten in Prozent. Mit größeren zugelassenen Spannweiten erhöht sich die Unsicherheit bei starken Blickwinkelveränderungen. Dass die Kurven am Ende bei 85° nochmals ansteigen ist auf die numerische Unsicherheit, bedingt durch die geringe Anzahl an gefundenen Zuordnungen, zurückzuführen.	58
6.8	Laufzeiten bei unterschiedlichen zugelassenen Spannweiten.	59
6.9	Simulation unterschiedlicher Kamerakonstanten. Die Rotationen ω und κ betragen jeweils 80 Grad. (a) Zeigt eine affine Abbildung ohne die perspektivischen Effekte. (b) zeigt eine perspektivische Abbildung mit einer simulierten normalen Perspektive. (c) zeigt eine perspektivische Abbildung mit einer simulierten kurzen Kamerakonstante, hier sind die perspektivischen Verkürzungen besonders gut sichtbar.	59
6.10	Ergebnisse 60 und 70 Grad bei einer simulierten kurzen Kamerakonstante, was sich im transformierten Bild in einer stärkeren perspektivischen Verkürzung äußert.	60
6.11	Die Auswirkungen von unterschiedlichen perspektivischen Deformationen bedingt durch unterschiedliche Kamerakonstanten.	61
6.12	Unterschiedliche Radiometrien haben keinen Einfluss auf die zugrunde liegende Struktur. $\omega = 50^\circ, P = 1.5$	62
6.13	Unterschiedliche Teilbereiche, die jeweils Teilmengen des anderen Bildes darstellen.	63
6.14	Anzahl Zuordnungen bei unterschiedlichen Teilbereichen.	66
6.15	Richtige Zuordnungen bei Teilbereichen in Prozent.	67
6.16	Unterschiedliche Eckpunkte in beiden Bildern. Diese können zu fehlerhaften Punktkorrespondenzen führen.	68
6.17	Richtige Zuordnungen in Abhängigkeit zu dem Anteil an fehlerhaften Punkten im Bild bei unterschiedlichen Blickwinkeln.	69
6.18	Einfluss des Reduktionswertes der Doppelverhältnisse auf die Rechenzeiten.	70
6.19	Einfluss des Reduktionswertes der Doppelverhältnisse auf die Anzahl und Richtigkeit der Punktzuordnungen.	71
6.20	2000 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.	72

Abbildungsverzeichnis

6.21	1000 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten.	72
6.22	500 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten. .	73
6.23	250 Pixel breites Bild: Anzahl Zuordnungen (oben), richtige Zuordnungen [%] (unten), jeweils bei unterschiedlichen I_1 und I_2 Spannweiten. .	73
6.24	Ticket-Automat aus zwei unterschiedlichen Blickwinkeln aufgenommen. Die perspektivischen Effekte sind an der Schrift gut zu sehen. Auch die farblichen Unterschiede in beiden Bildern sind gut sichtbar. 17 Zuordnungen.	75
6.25	Werbetafel eines Brezelstandes aus zwei unterschiedlichen Blickwinkeln fotografiert. 32 Zuordnungen, davon eine Fehlzuordnung.	75
6.26	Weitere Realbeispiele.	76
6.27	Testbilder zum testen mehrerer Ebenen. Die Kameraposition und die Positionen der Ebenen wurden verändert	76
6.28	Zuordnungen bei unterschiedlichen Ebenen. Jede grüne Linie entspricht einer kompletten Transformation.	77
6.29	Bildüberlagerung der aufeinander transformierten Ebenen. Links für identische Punkte aus Ebene <i>EINS</i> , rechts für Punkte aus Ebene <i>ZWEI</i> .	78

Tabellenverzeichnis

2.1 Eine Übersicht über die unterschiedlichen Typen von Geometrien in Abhängigkeit der verwendeten geometrischen Transformationen und der unterschiedlichen invarianten Eigenschaften (nach [Mikhail *et al.*, 2004] sowie [Faugeras & Luong, 2004], verändert). Im oberen Bereich der Tabelle sind die jeweiligen verwendeten Transformationen je Geometrie dargestellt. Im unteren Bereich sind die mit den jeweiligen Geometrien verbundenen invarianten Eigenschaften aufgelistet. Es ist sofort ersichtlich, dass mit jeder hinzukommenden Transformation invariante Eigenschaften verloren gehen. Bei der projektiven Geometrie bleiben lediglich Inzidenzen, Doppelverhältnisse, gerade Linien sowie Konnektivitäten als projektiv invariante Eigenschaften erhalten. Diese Invariante sind auch für alle anderen aufgeführten Geometrien und deren Transformationen gültig. 17

Literaturverzeichnis

- Agarwal, S, Snavely, N, Simon, I, Seitz, SM, & Szeliski, R. 2009. Building Rome in a day. *Pages 72–79 of: Computer Vision, 2009 IEEE 12th International Conference on.* IEEE.
- Agarwal, Sameer, Furukawa, Yasutaka, Snavely, Noah, Simon, Ian, Curless, Brian, Seitz, Steven M., & Szeliski, Richard. 2011. Building Rome in a Day. *Commun. ACM*, **54**(10), 105–112.
- Alahi, Alexandre, Ortiz, Raphael, & Vandergheynst, Pierre. 2012. Freak: Fast retina keypoint. *Pages 510–517 of: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on.* IEEE.
- Albertz, Jörg, & Wiggenhagen, Manfred. 2009. *Taschenbuch zur Photogrammetrie und Fernerkundung.* Wichmann.
- Allaire, S., Kim, J.J., Breen, S.L., Jaffray, D.A., & Pekar, V. 2008 (June). Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis. *Pages 1–8 of: Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on.*
- Araabi, B.N., Kehtarnavaz, N., Yearly, M., Hillman, G., & Wursig, B. 2002. Locating an affine/projective invariant identifier patch on an image. *Pages 121–125 of: Fifth IEEE Southwest Symposium on Image Analysis and Interpretation, 2002. Proceedings.*
- Bansal, Mayank, Daniilidis, Kostas, & Sawhney, Harpreet. 2012. *Ultra-wide Baseline Facade Matching for Geo-localization.* Vol. 7583. Springer Berlin Heidelberg. Pages 175–186.
- Barrett, Eamon B, Payton, Paul M, Haag, Nils N, & Brill, Michael H. 1991. General methods for determining projective invariants in imagery. *CVGIP: Image Understanding*, **53**(1), 46–65.
- Bartelsen, Jan, & Mayer, Helmut. 2010. Orientation of Image Sequences Acquired from UAVs and with GPS Cameras. *Surveying and Land Information Science*, **70**(3), 151–159.
- Bartelsen, Jan, Mayer, Helmut, Hirschmüller, Heiko, Kuhn, Andreas, & Michelini, Mario. 2012. Orientation and dense reconstruction from unordered wide baseline image sets. *Photogrammetrie-Fernerkundung-Geoinformation*, **2012**(4), 421–432.

- Bay, Herbert, Tuytelaars, Tinne, & Van Gool, Luc. 2006. *SURF: Speeded Up Robust Features*. Lecture Notes in Computer Science, vol. 3951. Springer Berlin Heidelberg. Pages 404–417.
- Bodenmueller, Tim. 2009. *Streaming surface reconstruction from real time 3D-measurements*. Ph.D. thesis, Technical University Munich.
- Branca, A., Stella, E., & Distanto, A. 1998 (Sep). Autonomous navigation of underwater vehicles. *Pages 61–65 vol.1 of: OCEANS '98 Conference Proceedings*, vol. 1.
- Bronstein, Alexander M., Bronstein, Michael M., Guibas, Leonidas J., & Ovsjanikov, Maks. 2011. Shape Google: Geometric Words and Expressions for Invariant Shape Retrieval. *ACM Trans. Graph.*, **30**(1), 1:1–1:20.
- Bryner, D., Klassen, E., Le, Huiling, & Srivastava, A. 2014. 2D Affine and Projective Shape Analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **36**(5), 998–1011.
- Bulatov, Dimitri. 2011. *Textured 3D reconstruction of urban terrain from UAV-borne video sequences*. Ph.D. thesis.
- Canny, John. 1986. A Computational Approach to Edge Detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **PAMI-8**(6), 679–698.
- Carlsson, Stefan and. 1998. Geometric structure and view invariant recognition. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **356**(1740), 1233–1250.
- Chotikakamthorn, N., Pantuwong, N., & Yawai, Wiyada. 2005 (Sept). Projective-invariant digital image watermarking technique using four co-planar feature points. *Pages 1–1005–8 of: Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 1.
- Coelho, Christopher, Heller, Aaron, Mundy, Joseph L., Forsyth, David A., & Zisserman, Andrew. 1992. *Geometric Invariance in Computer Vision*. Cambridge, MA, USA: MIT Press.
- Criminisi, Antonio. 2001. *Accurate visual metrology from single and multiple uncalibrated images*. Springer Science & Business Media.
- Delaunay, Boris. 1934. Sur la sphere vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvoennyka Nauk*, **7**(793-800), 1–2.
- Deseilligny, M Pierrot, & Clery, I. 2011. Apero, an open source bundle adjustment software for automatic calibration and orientation of set of images. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **38**, 5.
- Dürer, Albrecht. 1525. *Underweysung der messung mit dem zirkel und richtscheyt. Linien, Ebenen und Ganzen Corporen (Nuremberg)*.

- Fan, Xin, Luo, Zhongxuan, Zhang, Jieli, Zhou, Xinchun, Jia, Qi, & Luo, Daiyun. 2014. Characteristic Number: Theory and Its Application to Shape Analysis. *Axioms*, **3**(2), 202–221.
- Fan, Xin, Wang, Hao, Luo, Zhongxuan, Li, Yuntao, Hu, Wenyu, & Luo, Daiyun. 2015. Fiducial Facial Point Extraction Using a Novel Projective Invariant. *IEEE Transactions on Image Processing*, **24**(3), 1164–1177.
- Faugeras, Olivier, & Luong, Quang-Tuan. 2004. *The geometry of multiple images: the laws that govern the formation of multiple images of a scene and some of their applications*. MIT press.
- Fischler, Martin A., & Bolles, Robert C. 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, **24**(6), 381–395.
- Förstner, W., Dickscheid, T., & Schindler, F. 2009 (Sept). Detecting interpretable and accurate scale-invariant keypoints. *Pages 2256–2263 of: Computer Vision, 2009 IEEE 12th International Conference on*.
- Förstner, Wolfgang, & Gülch, Eberhard. 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. *Pages 281–305 of: Proc. ISPRS intercommission conference on fast processing of photogrammetric data*.
- Forsyth, D., Mundy, J.L., Zisserman, A., & Brown, C.M. 1990 (Dec). Invariance—a new framework for vision. *Pages 598–605 of: Third International Conference on Computer Vision, 1990. Proceedings*.
- Forsyth, D., Mundy, J.L., Zisserman, A., Coelho, C., Heller, A., & Rothwell, C. 1991. Invariant descriptors for 3D object recognition and pose. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **13**(10), 971–991.
- Frahm, Jan-Michael, Fite-Georgel, Pierre, Gallup, David, Johnson, Tim, Raguram, Rahul, Wu, Changchang, Jen, Yi-Hung, Dunn, Enrique, Clipp, Brian, Lazebnik, Svetlana, et al. . 2010. Building Rome on a cloudless day. *Pages 368–381 of: Computer Vision—ECCV 2010*. Springer.
- Gevers, Theo, & Smeulders, Arnold W. M. 1994. Efficient recognition of 3D rigid solid objects from 2D projective images based on projective invariant descriptions. *Pages 296–302 of: ISPRS Commission III Symposium: Spatial Information from Digital Photogrammetry and Computer Vision*, vol. 2357.
- Goodall, Colin R., & Mardia, Kanti V. 1999. Projective Shape Analysis. *Journal of Computational and Graphical Statistics*, **8**(2), 143–168.
- Gupta, R., Patil, H., & Mittal, A. 2010 (June). Robust order-based methods for feature description. *Pages 334–341 of: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010*.

- Gupta, Raj, & Mittal, Anurag. 2008. *SMD: A Locally Stable Monotonic Change Invariant Feature Descriptor*. Vol. 5303. Springer Berlin Heidelberg. Pages 265–277.
- Harris, Chris, & Stephens, Mike. 1988. A combined corner and edge detector. *Page 50 of: Alvey vision conference*, vol. 15. Manchester, UK.
- Hartley, R.I. 1994. Projective reconstruction and invariants from multiple images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **16**(10), 1036–1041.
- Heine, Heinrich. 1844. Doctrin. *Vorwärts! Pariser Deutsche Zeitschrift*, **48**(July), 1.
- Heinly, Jared, Dunn, Enrique, & Frahm, Jan-Michael. 2014. *Correcting for Duplicate Scene Structure in Sparse 3D Reconstruction*. Vol. 8692. Springer International Publishing. Pages 780–795.
- Hinz, Stefan. 2003. *Automatische Extraktion urbaner Straßennetze aus Luftbildern*. Ph.D. thesis, Technische Universität München, Universitätsbibliothek.
- Hirschmuller, H. 2008. Stereo Processing by Semiglobal Matching and Mutual Information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **30**(2), 328–341.
- Jia, Qi, Fan, Xin, Luo, Zhongxuan, Liu, Yu, & Guo, He. 2014. A new geometric descriptor for symbols with affine deformations. *Pattern Recognition Letters*, **40**(0), 128–135.
- Kliot, Michael, & Rivlin, Ehud. 1998. *Invariant-based shape retrieval in pictorial databases*. Vol. 1406. Springer Berlin Heidelberg. Pages 491–507.
- Koser, K., & Koch, R. 2007 (Oct). Perspectively Invariant Normal Features. *Pages 1–8 of: Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*.
- Kottman, Michal. 2011. Planar object detection using local feature descriptors. *Association of Computing Machinery bulletin*, **3**(2), 59–63.
- Kottman, Michal. 2013. Improving Binary Feature Descriptors Using Spatial Structure. *Information Sciences & Technologies: Bulletin of the ACM Slovakia*, **5**(2).
- Leutenegger, S., Chli, M., & Siegwart, R.Y. 2011 (Nov). BRISK: Binary Robust invariant scalable keypoints. *Pages 2548–2555 of: Computer Vision (ICCV), 2011 IEEE International Conference on*.
- Li, Linlin, & Tan, C.L. 2010. Recognizing Planar Symbols with Severe Perspective Deformation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **32**(4), 755–762.
- Li, Wen-Jing, & Lee, Tong. 2002. Invariant feature matching by Hopfield-type neural network. *Pages 2743–2748 of: Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on*, vol. 3.

- Li, Wen-Jing, Lee, Tong, & Tsui, Hung-Tat. 2002. Automatic feature matching using coplanar projective invariants for object recognition. *Pages 23–25 of: Proceedings of the 5th Asian Conference on Computer Vision (ACCV), Melbourne, Australia.*
- Lin, Wei-Yang. 2006. *Robust geometrically invariant features for 2d shape matching and 3d face recognition*. Ph.D. thesis, University Of Wisconsin–Madison.
- Lourakis, Manolis IA, & Argyros, Antonis A. 2009. SBA: A software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software (TOMS)*, **36**(1), 2.
- Lourakis, M.I.A, Halkidis, S.T, & Orphanoudakis, S.C. 2000. Matching disparate views of planar surfaces using projective invariants. *Image and Vision Computing*, **18**(9), 673 – 683.
- Lowe, David G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60**(2), 91–110.
- Lowe, D.G. 1999. Object recognition from local scale-invariant features. *Pages 1150–1157 vol.2 of: Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2.
- Lu, Xiaoye, & Manduchi, R. 2004 (June). Wide baseline feature matching using the cross-epipolar ordering constraint. *Pages I–16–I–23 Vol.1 of: Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1.
- Luhmann, T. 2010. *Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen*. Wichmann.
- Luo, Zhongxuan, Luo, Daiyun, Fan, Xin, Zhou, Xinchun, & Jia, Qi. 2013 (Sept). A shape descriptor based on new projective invariants. *Pages 2862–2866 of: Image Processing (ICIP), 2013 20th IEEE International Conference on*.
- Matas, J, Chum, O, Urban, M, & Pajdla, T. 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, **22**(10), 761 – 767. *British Machine Vision Computing 2002*.
- Mayer, Helmut, Bartelsen, Jan, Hirschmüller, Heiko, & Kuhn, Andreas. 2012. Dense 3d reconstruction from wide baseline image sets. *Pages 285–304 of: Outdoor and Large-Scale Real-World Scene Analysis*. Springer.
- McManus, Colin, Upcroft, Ben, & Newmann, Paul. 2014 (July). Scene signatures : localised and point-less features for localisation. *In: Robotics: Science and Systems X*.
- Meer, Peter, Ramakrishna, Sudhir, & Lenz, Reiner. 1994. *Correspondence of coplanar features through P2-invariant representations*. Vol. 825. Springer Berlin Heidelberg. Pages 473–492.

- Meer, Peter, Lenz, Reiner, & Ramakrishna, Sudhir. 1998. Efficient Invariant Representations. *International Journal of Computer Vision*, **26**(2), 137–152.
- Mikhail, Edward M, Bethel, James S. (James Samuel), McGlone, J. Chris, of Photogrammetry, American Society, Sensing, Remote, & Society, Imaging & Geospatial Information. 2004. *Manual of photogrammetry*. 5th ed edn. Bethesda, Md. : American Society of Photogrammetry and Remote Sensing. ÄSPRS, The Imaging & Geospatial Information Society".
- Mikolajczyk, K., & Schmid, C. 2001. Indexing based on scale invariant interest points. *Pages 525–531 vol.1 of: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1.
- Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., & Gool, L.Van. 2005. A Comparison of Affine Region Detectors. *International Journal of Computer Vision*, **65**(1-2), 43–72.
- Mikolajczyk, Krystian, & Schmid, Cordelia. 2002. *An Affine Invariant Interest Point Detector*. Vol. 2350. Springer Berlin Heidelberg. Pages 128–142.
- Moisan, Lionel, Moulon, Pierre, & Monasse, Pascal. 2012. Automatic homographic registration of a pair of images, with a contrario elimination of outliers. *Image Processing On Line*, **2**, 56–73.
- Moravec, HP. 1977. Towards automatic visual obstacle avoidance. *Page 584 of: Proceedings of the international joint conference on artificial intelligence (IJCAI)*, vol. 2.
- Mundy, J.L., & Zisserman, A. 1992. *Geometric Invariance in Computer Vision*. Artificial intelligence. MIT Press.
- Nister, D. 2003 (June). An efficient solution to the five-point relative pose problem. *Pages II–195–202 vol.2 of: Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 2.
- Nister, D. 2004. An efficient solution to the five-point relative pose problem. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **26**(6), 756–770.
- Pach, János, & Sharir, Micha. 2004. Geometric incidences. *Contemporary Mathematics*, **342**, 185–224.
- Patterson, Boyd C. 1928. The differential invariants of inversive geometry. *American Journal of Mathematics*, 553–568.
- Planck, Max. 1948. *Wissenschaftliche Selbstbiographie*. Johann Ambrosius Barth Verlag, Leipzig.
- Prasad, Dilip K. 2013. *Geometric primitive feature extraction - concepts, algorithms, and applications*. Ph.D. thesis.

- Putjarupong, Parichart, Pintavirooj, Chuchart, Withayachumnankul, Withawat, & Sangworasil, Manas. 2004. Image registration exploiting five-point coplanar perspective invariant and maximum-curvature point.
- Quan, Long. 1995. Invariants of six points and projective reconstruction from three uncalibrated images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(1), 34–46.
- Reich, M, Unger, J, Rottensteiner, F, & Heipke, C. 2013. On-Line Compatible Orientation of a Micro-Uav Based on Image Triplets. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1(2), 37–42.
- Ren, Yi, Bugeau, A., & Benois-Pineau, J. 2014 (Oct). Bag-of-bags of words irregular graph pyramids vs spatial pyramid matching for image retrieval. *Pages 1–6 of: Image Processing Theory, Tools and Applications (IPTA), 2014 4th International Conference on*.
- Rosten, Edward, & Drummond, Tom. 2005 (Oct). Fusing points and lines for high performance tracking. *Pages 1508–1515 Vol. 2 of: Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2.
- Rosten, Edward, & Drummond, Tom. 2006. *Machine Learning for High-Speed Corner Detection*. Vol. 3951. Springer Berlin Heidelberg. Pages 430–443.
- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. 2011 (Nov). ORB: An efficient alternative to SIFT or SURF. *Pages 2564–2571 of: Computer Vision (ICCV), 2011 IEEE International Conference on*.
- Schneider, D, Schwalbe, E, & Maas, H-G. 2009. Validation of geometric models for fisheye lenses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(3), 259–266.
- Sivic, J., & Zisserman, A. 2003 (Oct). Video Google: a text retrieval approach to object matching in videos. *Pages 1470–1477 vol.2 of: Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*.
- Snavely, Noah, Seitz, Steven M., & Szeliski, Richard. 2006. Photo Tourism: Exploring Photo Collections in 3D. *ACM Trans. Graph.*, 25(3), 835–846.
- Sobel, Irwin, & Feldman, Gary. 1968. A 3x3 isotropic gradient operator for image processing.
- Strecha, C., von Hansen, W., Van Gool, L., Fua, P., & Thoennessen, U. 2008 (June). On benchmarking camera calibration and multi-view stereo for high resolution imagery. *Pages 1–8 of: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*.
- Strunz, Günter. 1993. *Bildorientierung und Objektrekonstruktion mit Punkten, Linien und Flächen*. Munchen.

Literaturverzeichnis

- Tang, Feng, Lim, Suk Hwan, Chang, N.L., & Tao, Hai. 2009 (June). A novel feature descriptor invariant to complex brightness changes. *Pages 2631–2638 of: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on.*
- Tao, Yi, & Grosky, William I. 1998. Delaunay triangulation for image object indexing: a novel method for shape representation. vol. 3656.
- Tuytelaars, Tinne, & Mikolajczyk, Krystian. 2008. Local Invariant Feature Detectors: A Survey. *Found. Trends. Comput. Graph. Vis.*, **3**(3), 177–280.
- Tuytelaars, Tinne, & Van Gool, Luc. 2004. Matching Widely Separated Views Based on Affine Invariant Regions. *International Journal of Computer Vision*, **59**(1), 61–85.
- Ulrich, Markus. 2003. *Hierarchical Real-Time Recognition of Compound Objects in Images*. Dissertation, Technische Universität München, München.
- van Liere, R., & Mulder, J.D. 2003 (March). Optical tracking using projective invariant marker pattern properties. *Pages 191–198 of: Virtual Reality, 2003. Proceedings. IEEE.*
- Verdie, Yannick, Yi, Kwang Moo, Fua, Pascal, & Lepetit, Vincent. 2015. TILDE: A Temporally Invariant Learned DETector. *In: Proceedings of the Computer Vision and Pattern Recognition.*
- Verhagen, B., Timofte, R., & Van Gool, L. 2014 (March). Scale-invariant line descriptors for wide baseline matching. *Pages 493–500 of: Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on.*
- von Hundelshausen, F., & Sukthankar, R. 2012 (June). D-Nets: Beyond patch-based image descriptors. *Pages 2941–2948 of: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on.*
- Wang, Guang-Hui, Hu, Zhan-Yi, & Wu, Fu-Chao. 2004. Single view based measurement on space planes. *Journal of Computer Science and Technology*, **19**(3), 374–382.
- Wang, Zhenhua, Fan, Bin, & Wu, Fuchao. 2011 (Nov). Local Intensity Order Pattern for feature description. *Pages 603–610 of: Computer Vision (ICCV), 2011 IEEE International Conference on.*
- Wang, Zhiheng, Wu, Fuchao, & Hu, Zhanyi. 2009. MSLD: A robust descriptor for line matching. *Pattern Recognition*, **42**(5), 941 – 953.
- Weinmann, Martin. 2013. *Visual Features—From Early Concepts to Modern Computer Vision*. Springer London. Pages 1–34.
- Weiss, I. 1988 (Jun). Projective invariants of shapes. *Pages 291–297 of: Computer Vision and Pattern Recognition, 1988. Proceedings CVPR '88., Computer Society Conference on.*
- Weiss, Isaac. 1993. Geometric invariants and object recognition. *International Journal of Computer 11263on*, **10**(3), 207–231.

Literaturverzeichnis

- Yu, Guoshen, & Morel, J.-M. 2009 (April). A fully affine invariant image comparison method. *Pages 1597–1600 of: Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on.*
- Zhang, Dengsheng, & Lu, Guojun. 2002. Generic Fourier descriptor for shape-based image retrieval. *Pages 425–428 vol.1 of: Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on, vol. 1.*
- Zhang, Dengsheng, & Lu, Guojun. 2004. Review of shape representation and description techniques. *Pattern recognition, 37(1), 1–19.*