

**GEOMETRISCHE UND STOCHASTISCHE MODELLE  
ZUR VERARBEITUNG VON 3D-KAMERADATEN  
AM BEISPIEL MENSCHLICHER BEWEGUNGSANALYSEN**

Dissertation zur Erlangung des akademischen Grades  
Doktor-Ingenieur (Dr.-Ing.)

vorgelegt von

**Dipl.-Ing. Patrick Westfeld**

Dresden im Mai 2012

Das vorliegende Exemplar entspricht einer leicht überarbeiteten, ergänzten und von den Gutachtern genehmigten Version der Originaldissertation *Geometrische und stochastische Modelle zur Verarbeitung von 3D-Kameradaten am Beispiel menschlicher Bewegungsanalysen*.

Patrick Westfeld

Dresden, 10.05.2012

1. Gutachter: Herr Prof. Dr. sc. techn. habil. Hans-Gerd Maas, Technische Universität Dresden
2. Gutachter: Herr Prof. Dr. rer. nat. Ralf Reulke, Humboldt-Universität zu Berlin

Tag der Einreichung: 15. November 2011

Tag der Verteidigung: 8. Mai 2012

Für Lilo





## Zusammenfassung

Die dreidimensionale Erfassung der Form und Lage eines beliebigen Objekts durch die flexiblen Methoden und Verfahren der Photogrammetrie spielt für ein breites Spektrum technisch-industrieller und naturwissenschaftlicher Einsatzgebiete eine große Rolle. Die Anwendungsmöglichkeiten reichen von Messaufgaben im Automobil-, Maschinen- und Schiffbau über die Erstellung komplexer 3D-Modelle in Architektur, Archäologie und Denkmalpflege bis hin zu Bewegungsanalysen in Bereichen der Strömungsmesstechnik, Ballistik oder Medizin. In der Nahbereichsphotogrammetrie werden dabei verschiedene optische 3D-Messsysteme verwendet. Neben flächenhaften Halbleiterkameras im Einzel- oder Mehrbildverband kommen aktive Triangulationsverfahren zur Oberflächenmessung mit z. B. strukturiertem Licht oder Laserscanner-Systeme zum Einsatz.

3D-Kameras auf der Basis von Photomischdetektoren oder vergleichbaren Prinzipien erzeugen durch die Anwendung von Modulationstechniken zusätzlich zu einem Grauwertbild simultan ein Entfernungsbild. Als Einzelbildsensoren liefern sie ohne die Notwendigkeit einer stereoskopischen Zuordnung räumlich aufgelöste Oberflächendaten in Videorate. In der 3D-Bewegungsanalyse ergeben sich bezüglich der Komplexität und des Rechenaufwands erhebliche Erleichterungen. 3D-Kameras verbinden die Handlichkeit einer Digitalkamera mit dem Potential der dreidimensionalen Datenakquisition etablierter Oberflächenmesssysteme. Sie stellen trotz der noch vergleichsweise geringen räumlichen Auflösung als monosensorielles System zur Echtzeit-Tiefenbildakquisition eine interessante Alternative für Aufgabenstellungen der 3D-Bewegungsanalyse dar.

Der Einsatz einer 3D-Kamera als Messinstrument verlangt die Modellierung von Abweichungen zum idealen Abbildungsmodell; die Verarbeitung der erzeugten 3D-Kameradaten bedingt die zielgerichtete Adaption, Weiter- und Neuentwicklung von Verfahren der *Computer Vision* und Photogrammetrie. Am Beispiel der Untersuchung des zwischenmenschlichen Bewegungsverhaltens sind folglich die Entwicklung von Verfahren zur Sensorkalibrierung und zur 3D-Bewegungsanalyse die Schwerpunkte der Dissertation. Eine 3D-Kamera stellt aufgrund ihres inhärenten Designs und Messprinzips gleichzeitig Amplituden- und Entfernungsinformationen zur Verfügung, welche aus einem Messsignal rekonstruiert werden. Die simultane Einbeziehung aller 3D-Kamerainformationen in jeweils einen integrierten Ansatz ist eine logische Konsequenz und steht im Vordergrund der Verfahrensentwicklungen. Zum einen stützen sich die komplementären Eigenschaften der Beobachtungen durch die Herstellung des funktionalen Zusammenhangs der Messkanäle gegenseitig, wodurch Genauigkeits- und Zuverlässigkeitssteigerungen zu erwarten sind. Zum anderen gewährleistet das um eine Varianzkomponentenschätzung erweiterte stochastische Modell eine vollständige Ausnutzung des heterogenen Informationshaushalts.

Die entwickelte integrierte Bündelblockausgleichung ermöglicht die Bestimmung der exakten 3D-

Kamerageometrie sowie die Schätzung der distanzmessspezifischen Korrekturparameter zur Modellierung linearer, zyklischer und signalwegeeffektbedingter Fehleranteile einer 3D-Kamerastreckenmessung. Die integrierte Kalibrieroutine gleicht in beiden Informationskanälen gemessene Größen gemeinsam, unter der automatischen Schätzung optimaler Beobachtungsgewichte, aus. Die Methode basiert auf dem flexiblen Prinzip einer Selbstkalibrierung und benötigt keine Objektrauminformation, wodurch insbesondere die aufwendige Ermittlung von Referenzstrecken übergeordneter Genauigkeit entfällt. Die durchgeführten Genauigkeitsuntersuchungen bestätigen die Richtigkeit der aufgestellten funktionalen Zusammenhänge, zeigen aber auch Schwächen aufgrund noch nicht parametrisierter distanzmessspezifischer Fehler. Die Adaptivität und die modulare Implementierung des entwickelten mathematischen Modells gewährleisten aber eine zukünftige Erweiterung. Die Qualität der 3D-Neupunktkoordinaten kann nach einer Kalibrierung mit 5 mm angegeben werden. Für die durch eine Vielzahl von meist simultan auftretenden Rauschquellen beeinflusste Tiefenbildtechnologie ist diese Genauigkeitsangabe sehr vielversprechend, vor allem im Hinblick auf die Entwicklung von auf korrigierten 3D-Kameradaten aufbauenden Auswertalgorithmen.

*2,5D Least Squares Tracking (LST)* ist eine im Rahmen der Dissertation entwickelte integrierte spatiale und temporale Zuordnungsmethode zur Auswertung von 3D-Kamerabildsequenzen. Der Algorithmus basiert auf der in der Photogrammetrie bekannten Bildzuordnung nach der Methode der kleinsten Quadrate und bildet kleine Oberflächensegmente konsekutiver 3D-Kameradaten aufeinander ab. Die Abbildungsvorschrift wurde, aufbauend auf einer 2D-Affintransformation, an die Datenstruktur einer 3D-Kamera angepasst. Die geschlossen formulierte Parametrisierung verknüpft sowohl Grau- als auch Entfernungswerte in einem integrierten Modell. Neben den affinen Parametern zur Erfassung von Translations- und Rotationseffekten, modellieren die Maßstabs- sowie Neigungsparameter perspektivbedingte Größenänderungen des Bildausschnitts, verursacht durch Distanzänderungen in Aufnahmerichtung. Die Eingabedaten sind in einem Vorverarbeitungsschritt mit Hilfe der entwickelten Kalibrieroutine um ihre opto- und distanzmessspezifischen Fehler korrigiert sowie die gemessenen Schrägstrecken auf Horizontalabstände reduziert worden. *2,5D-LST* liefert als integrierter Ansatz vollständige 3D-Verschiebungsvektoren. Weiterhin können die aus der Fehlerrechnung resultierenden Genauigkeits- und Zuverlässigkeitsangaben als Entscheidungskriterien für die Integration in einer anwendungsspezifischen Verarbeitungskette Verwendung finden. Die Validierung des Verfahrens zeigte, dass die Einführung komplementärer Informationen eine genauere und zuverlässigere Lösung des Korrespondenzproblems bringt, vor allem bei schwierigen Kontrastverhältnissen in einem Kanal. Die Genauigkeit der direkt mit den Distanzkorrekturtermen verknüpften Maßstabs- und Neigungsparameter verbesserte sich deutlich. Darüber hinaus brachte die Erweiterung des geometrischen Modells insbesondere bei der Zuordnung natürlicher, nicht gänzlich ebener Oberflächensegmente signifikante Vorteile.

Die entwickelte flächenbasierte Methode zur Objektzuordnung und Objektverfolgung arbeitet auf der Grundlage berührungslos aufgenommener 3D-Kameradaten. Sie ist somit besonders für Aufgabenstellungen der 3D-Bewegungsanalyse geeignet, die den Mehraufwand einer multiokularen Experimentalanordnung und die Notwendigkeit einer Objektsignalisierung mit Zielmarken vermeiden

möchten. Das Potential des 3D-Kamerazuordnungsansatzes wurde an zwei Anwendungsszenarien der menschlichen Verhaltensforschung demonstriert. 2,5D-LST kam zur Bestimmung der interpersonalen Distanz und Körperorientierung im erziehungswissenschaftlichen Untersuchungsgebiet der Konfliktregulation befreundeter Kindespaare ebenso zum Einsatz wie zur Markierung und anschließenden Klassifizierung von Bewegungseinheiten sprachbegleitender Handgesten. Die Implementierung von 2,5D-LST in die vorgeschlagenen Verfahren ermöglichte eine automatische, effektive, objektive sowie zeitlich und räumlich hochaufgelöste Erhebung und Auswertung verhaltensrelevanter Daten.

Die vorliegende Dissertation schlägt die Verwendung einer neuartigen 3D-Tiefenbildkamera zur Erhebung menschlicher Verhaltensdaten vor. Sie präsentiert sowohl ein zur Datenaufbereitung entwickeltes Kalibrierwerkzeug als auch eine Methode zur berührungslosen Bestimmung dichter 3D-Bewegungsvektorfelder. Die Arbeit zeigt, dass die Methoden der Photogrammetrie auch für bewegungsanalytische Aufgabenstellungen auf dem bisher noch wenig erschlossenen Gebiet der Verhaltensforschung wertvolle Ergebnisse liefern können. Damit leistet sie einen Beitrag für die derzeitigen Bestrebungen in der automatisierten videographischen Erhebung von Körperbewegungen in dyadischen Interaktionen.



## Summary

The three-dimensional documentation of the form and location of any type of object using flexible photogrammetric methods and procedures plays a key role in a wide range of technical-industrial and scientific areas of application. Potential applications include measurement tasks in the automotive, machine building and ship building sectors, the compilation of complex 3D models in the fields of architecture, archaeology and monumental preservation and motion analyses in the fields of flow measurement technology, ballistics and medicine. In the case of close-range photogrammetry a variety of optical 3D measurement systems are used. Area sensor cameras arranged in single or multi-image configurations are used besides active triangulation procedures for surface measurement (e.g. using structured light or laser scanner systems).

The use of modulation techniques enables 3D cameras based on photomix detectors or similar principles to simultaneously produce both a grey value image and a range image. Functioning as single image sensors, they deliver spatially resolved surface data at video rate without the need for stereoscopic image matching. In the case of 3D motion analyses in particular, this leads to considerable reductions in complexity and computing time. 3D cameras combine the practicality of a digital camera with the 3D data acquisition potential of conventional surface measurement systems. Despite the relatively low spatial resolution currently achievable, as a monosensory real-time depth image acquisition system they represent an interesting alternative in the field of 3D motion analysis.

The use of 3D cameras as measuring instruments requires the modelling of deviations from the ideal projection model, and indeed the processing of the 3D camera data generated requires the targeted adaptation, development and further development of procedures in the fields of computer graphics and photogrammetry. This Ph. D. thesis therefore focuses on the development of methods of sensor calibration and 3D motion analysis in the context of investigations into inter-human motion behaviour. As a result of its intrinsic design and measurement principle, a 3D camera simultaneously provides amplitude and range data reconstructed from a measurement signal. The simultaneous integration of all data obtained using a 3D camera into an integrated approach is a logical consequence and represents the focus of current procedural development. On the one hand, the complementary characteristics of the observations made support each other due to the creation of a functional context for the measurement channels, with is to be expected to lead to increases in accuracy and reliability. On the other, the expansion of the stochastic model to include variance component estimation ensures that the heterogeneous information pool is fully exploited.

The integrated bundle adjustment developed facilitates the definition of precise 3D camera geometry and the estimation of range-measurement-specific correction parameters required for the modelling of the linear, cyclical and latency defectives of a distance measurement made using a 3D camera.

The integrated calibration routine jointly adjusts appropriate dimensions across both information channels, and also automatically estimates optimum observation weights. The method is based on the same flexible principle used in self-calibration, does not require spatial object data and therefore foregoes the time-consuming determination of reference distances with superior accuracy. The accuracy analyses carried out confirm the correctness of the proposed functional contexts, but nevertheless exhibit weaknesses in the form of non-parameterized range-measurement-specific errors. This notwithstanding, the future expansion of the mathematical model developed is guaranteed due to its adaptivity and modular implementation. The accuracy of a new 3D point coordinate can be set at 5 mm further to calibration. In the case of depth imaging technology – which is influenced by a range of usually simultaneously occurring noise sources – this level of accuracy is very promising, especially in terms of the development of evaluation algorithms based on corrected 3D camera data.

*2.5D Least Squares Tracking (LST)* is an integrated spatial and temporal matching method developed within the framework of this Ph. D. thesis for the purpose of evaluating 3D camera image sequences. The algorithm is based on the least squares image matching method already established in photogrammetry, and maps small surface segments of consecutive 3D camera data sets on top of one another. The mapping rule has been adapted to the data structure of a 3D camera on the basis of a 2D affine transformation. The closed parameterization combines both grey values and range values in an integrated model. In addition to the affine parameters used to include translation and rotation effects, the scale and inclination parameters model perspective-related deviations caused by distance changes in the line of sight. A pre-processing phase sees the calibration routine developed used to correct optical and distance-related measurement specific errors in input data and measured slope distances reduced to horizontal distances. *2.5D LST* is an integrated approach, and therefore delivers fully three-dimensional displacement vectors. In addition, the accuracy and reliability data generated by error calculation can be used as decision criteria for integration into an application-specific processing chain. Process validation showed that the integration of complementary data leads to a more accurate, reliable solution to the correspondence problem, especially in the case of difficult contrast ratios within a channel. The accuracy of scale and inclination parameters directly linked to distance correction terms improved dramatically. In addition, the expansion of the geometric model led to significant benefits, and in particular for the matching of natural, not entirely planar surface segments.

The area-based object matching and object tracking method developed functions on the basis of 3D camera data gathered without object contact. It is therefore particularly suited to 3D motion analysis tasks in which the extra effort involved in multi-ocular experimental settings and the necessity of object signalling using target marks are to be avoided. The potential of the 3D camera matching approach has been demonstrated in two application scenarios in the field of research into human behaviour. As in the case of the use of *2.5D LST* to mark and then classify hand gestures accompanying verbal communication, the implementation of *2.5D LST* in the proposed procedures for the determination of interpersonal distance and body orientation within the framework of pedagogical research into conflict regulation between pairs of child-age friends facilitates the automatic, effective, objecti-

ve and high-resolution (from both a temporal and spatial perspective) acquisition and evaluation of data with relevance to behaviour.

This Ph.D. thesis proposes the use of a novel 3D range imaging camera to gather data on human behaviour, and presents both a calibration tool developed for data processing purposes and a method for the contact-free determination of dense 3D motion vector fields. It therefore makes a contribution to current efforts in the field of the automated videographic documentation of bodily motion within the framework of dyadic interaction, and shows that photogrammetric methods can also deliver valuable results within the framework of motion evaluation tasks in the as-yet relatively untapped field of behavioural research.





# Inhaltsverzeichnis

I	Einleitung	17
1	<b>Motivation – 19</b>	
2	<b>Zielsetzung – 25</b>	
3	<b>Gliederung der Arbeit – 27</b>	
II	Theoretische Grundlagen	29
4	<b>Mathematische Grundlagen – 31</b>	
4.1	Ebene Koordinatentransformation – 31	
4.2	Räumliche Koordinatentransformation – 32	
4.3	Räumliche Drehung – 33	
4.4	Ausgleichsrechnung – 34	
4.5	Statistische Qualitätsmaße – 40	
4.6	Varianzkomponentenschätzung – 44	
4.7	Random Sample Consensus – 48	
5	<b>Photogrammetrische und geodätische Grundlagen – 53</b>	
5.1	Abbildungsmodell – 53	
5.2	Objektrekonstruktion – 57	
5.3	Einzelbildorientierung – 57	
5.4	Mehrbildtriangulation – 60	
6	<b>Verfahren zur 3D-Objekterfassung – 65</b>	
6.1	Übersicht – 65	
6.2	Triangulation – 66	
6.3	Laufzeitmessung – 72	
6.4	Diskussion – 75	
7	<b>Bewegungsanalyse – 77</b>	
7.1	Übersicht – 79	
7.2	Bildzuordnung – 83	
7.3	Probleme bei der Bildzuordnung – 86	
7.4	Kreuzkorrelation – 89	

- 7.5 Methoden des Optischen Flusses – 91
- 7.6 Least Squares Matching – 94
- 7.7 Diskussion – 100

### III Sensormodellierung

101

- 8 Distanzmessende Kamera – 103**
  - 8.1 Systemkomponenten – 105
  - 8.2 Messprinzip – 108
  - 8.3 Geometrisches Modell – 110
  - 8.4 Anwendungsbeispiele – 112
  - 8.5 Diskussion – 114
- 9 Fehlereinflüsse und deren Korrektur – 117**
  - 9.1 Temperatur – 117
  - 9.2 Hintergrundstrahlung – 118
  - 9.3 Mehrwegeempfang – 119
  - 9.4 Fixed Pattern Noise – 120
  - 9.5 Distanzabhängige Fehlereinflüsse – 121
  - 9.6 Signalwegeeffektbedingte Fehler – 123
  - 9.7 Reflektivität – 123
  - 9.8 Bewegungsartefakte – 124
  - 9.9 Diskussion – 125
- 10 Kalibrierstrategien – 127**
  - 10.1 Photogrammetrische Kalibrierung – 127
  - 10.2 Sequentielle photogrammetrische und Distanzkalibrierung – 128
  - 10.3 Simultane photogrammetrische und Distanzkalibrierung – 129
  - 10.4 Diskussion – 131
- 11 Integrierte Bündelblockausgleichung – 133**
  - 11.1 Ausgleichungsansatz – 134
  - 11.2 Bereitstellung der Beobachtungen – 136
  - 11.3 Erläuterung der Unbekannten und Beschaffung von Näherungswerten – 144
  - 11.4 Zusätzliche Bedingungen – 147
  - 11.5 Funktionales Modell – 149
  - 11.6 Stochastisches Modell – 153
  - 11.7 Lösung der Ausgleichungsaufgabe – 155
  - 11.8 Validierung der integrierten Bündelblockausgleichung – 158
  - 11.9 Diskussion – 170

---

IV	Bewegung	173
12	<b>Verfahren zur Erfassung menschlicher Bewegungen – 175</b>	
12.1	Bewegungsinformation durch interaktive Videografie – 176	
12.2	Bewegungsinformation durch automatische Bildsequenzanalyse – 177	
12.3	Trackingsysteme – 182	
12.4	Diskussion – 184	
13	<b>Analyse von 3D-Kamerabildsequenzen – 185</b>	
13.1	Auswertung des Amplituden- oder Entfernungskanals – 186	
13.2	Sequentielle Auswertung des Amplituden- und Entfernungskanals – 188	
13.3	Simultane Auswertung des Amplituden- und Entfernungskanals – 191	
13.4	Diskussion – 193	
14	<b>Integriertes Least Squares Tracking – 195</b>	
14.1	Übersicht – 195	
14.2	Funktionales Modell – 196	
14.3	Stochastisches Modell – 205	
14.4	Lösung der Ausgleichungsaufgabe – 206	
14.5	Validierung des integrierten Least Squares Tracking – 211	
14.6	Diskussion – 222	
15	<b>Anwendungen der nonverbalen Interaktionsforschung – 225</b>	
15.1	Konfliktregulation bei befreundeten Kindespaaren – 225	
15.2	Gestenerkennung – 228	
15.3	Diskussion – 233	
V	Zusammenfassung und Ausblick	235
16	<b>Zusammenfassung – 237</b>	
17	<b>Ausblick – 241</b>	
	<b>Literaturverzeichnis – 247</b>	
	<b>Abbildungsverzeichnis – 271</b>	
	<b>Tabellenverzeichnis – 275</b>	
	<b>Abkürzungsverzeichnis – 277</b>	







## Motivation

Die Untersuchung des zwischenmenschlichen Verhaltens besitzt einen großen Stellenwert in den Erziehungswissenschaften und in der Psychologie. Dabei wird häufig das Interaktionsverhalten zweier Personen in so genannten dyadischen Interaktionen, z. B. zwischen Mutter-Kind (Aschersleben & Henning 2008), Patient-Therapeut (Lausberg u. a. 2007), Schüler-Lehrer (Alisch 2010) und Kind-Kind (Altmann 2010) fokussiert (Abbildung 1.1). Altmann u. a. (2007) erklären in diesem Zusammenhang, dass das im wesentlichen Maße implizit menschliche Interaktionen regulierende nonverbale Verhalten der interagierenden Personen zumeist harmonisch aufeinander abgestimmt ist. Dies äußert sich bei gesprächsbegleitenden kommunikativen und selbstregulierenden Bewegungsaktivitäten in bspw. synchronen Positionswechseln, synchron auftretenden Selbstberührungen, dem Nachahmen von Mimik und Gestik oder in der Spiegelung der Körperhaltung. Mögliche Abweichungen von dieser interaktiven Rhythmizität und Synchronizität geben dem Verhaltensforscher Aufschluss über eine fehlende AffektAbstimmung und gelten als Indikator für das Vorliegen einer gestörten zwischenmenschlichen Beziehung oder gar psychischen Erkrankung. Die Dissertationsschrift von Altmann (2012) gibt einen sehr ausführlichen Forschungsüberblick aus erziehungswissenschaftlicher Sicht.



(a)



(b)



(c)

**Abb. 1.1:** Zwischenmenschliches Verhalten: (a) Eine sichere Bindung zwischen Mutter und Kind ist essentiell für eine gesunde seelische Entwicklung des Babys (Aschersleben & Henning 2008). (b) Die Interaktion zwischen Patient und Psychotherapeut steht in einem engen Zusammenhang mit dem Therapieerfolg. (c) Die Analyse von Konflikten und deren Regulation in Kinderfreundschaften ermöglicht, Aussagen über die soziale Entwicklung eines Kindes zu treffen.

Die Erhebung von Ort und Zeitpunkt des Auftretens eines relevanten Ereignisses sowie die Klassifikation und anschließende Interpretation des Phänomens geschehen mit Kodierverfahren. Solche Verfahren dokumentieren das sichtbare Verhalten und schaffen durch die Definition und Beschreibung von Bewegungseinheiten (*action units*) objektive sowie mess- und vergleichbare Werte. Be-

deutende Kodierverfahren sind z. B. das *Facial Action Coding System (FACS)* zur Kodierung von Bewegungen im Gesicht (Ekman u. a. 1978), das Berner System zur Erfassung von Körperhaltungen (Frey u. a. 1981) und das *Neuropsychological Gesture Coding System (NEUROGES)* zur Dokumentation sprachbegleitender Handgesten.

Die Erfassung eines Interaktionsverlaufs beruht zumeist auf Vor-Ort-Beobachtungen, teilweise mit unterstützender Videoaufzeichnung, und anschließender manueller Kodierung der Bewegungseinheiten entsprechend des gewählten Erhebungsverfahrens (Abbildung 1.2). Der Wunsch nach einer Automatisierung wurde dabei bereits früh geäußert: Webb u. a. (1975) kritisieren den indirekten Einfluss des Beobachters auf das Verhalten der Interagierenden; Kempton (1980) betrachtet den erhöhten Zeit- und Personalaufwand einer manuellen Kodierung, vor allem bei zeitlich hochaufgelösten Studien oder der unabhängigen Einschätzung durch mehrere Kodierer.

Dennoch wurden die teils bis heute etablierten Papier- und Bleistift-Verfahren nur sehr zögerlich zu Gunsten einer Steigerung der Effektivität durch computergestützte Videografieauswertungen ersetzt. Die vom Max-Planck-Institut für Psycholinguistik (Nijmegen, NL) kostenlos zur Verfügung gestellte Software *ELAN* ermöglicht z. B. eine professionelle Annotation von Audio- oder Videodaten am Rechner. Diese wird bspw. von Lausberg und Kryger (2011) routinemäßig genutzt, um, entsprechend dem *NEUROGES*, Veränderungen des gestischen Verhaltens eines leicht depressiven Patienten zu Beginn und am Ende einer erfolgreichen Psychotherapie zu analysieren (Abbildung 1.3).

Fällt zusätzlich der Arbeitsschritt der interaktiven Annotation weg, kann von einer automatisierten und (damit einhergehenden) objektiven Erhebung verhaltensrelevanter Körperbewegungen gesprochen werden. Für die Interaktionsforschung benennt Altmann (2012) zwei bedeutende Verfahren der automatisierten Kodierung: Beim *Motion-Capture*-Verfahren werden kommerziell verfügbare Systeme zur vollständigen Erfassung der Bewegungsabläufe von Probanden eingesetzt. Boker u. a. (2005) untersuchen bspw. den Grad der Symmetrie der Körperhaltungen bei tanzenden, gemischtgeschlechtlichen Paaren, welche, mit jeweils acht Bewegungssensoren präpariert, in einem induzierten Magnetfeld verfolgt wurden. In (Kawaguchi u. a. 2010) wird das Verhalten von mit passiven Zielmarken signalisierten autistischen Kindern in Spielsituationen mit einem Mehrkamerasystem aufgenommen und analysiert. Die Analyse der Bewegungsenergie (*Motion Energy Analysis, MEA*) als zweites wichtiges Verfahren basiert aus photogrammetrischer Sicht auf einer einfachen Differenzbildung konsekutiver Bilder einer Sequenz und der schwellenwertbasierten Detektion von Veränderungen unterliegenden Sensorbereichen. Grammer u. a. (2003) untersuchen so Tanzbewegungen im Zusammenhang mit der Attraktivität der Tänzerin, Ramseyer (2010) quantifiziert die nonverbale Synchronisation in Psychotherapien (Abbildung 1.5) und Altmann (2010) erhebt die Nonverbalität in einem Experiment zur Kooperation von Schülern.

Die vorgestellten und auf dem Gebiet der Erziehungswissenschaften und der Psychologie etablierten interaktiven und automatisierten Verfahren zur Erhebung von Körperbewegungen erhöhen zwar die Effektivität und die Objektivität des Auswertungsprozesses, gehen aber mit wesentlichen Nachteilen einher: Die interaktiv mit einer Annotationssoftware durchgeführte Analyse erleichtert die Kodierung eines Videos, ist jedoch vor allem in datenintensiven Studien noch immer zeit- und perso-



## Facial Action Coding System: Score Sheet # 1

Designed by Paul Ekman and Wallace V. Friesen

## LOWER FACE

I. Initial Scoring \_\_\_\_\_  
 II. Omission Check \_\_\_\_\_  
 III. Reorganized Scoring \_\_\_\_\_  
 IV. Co-Occurrence Rule Check  
 AUs Numeric Order \_\_\_\_\_  
 Rules: \_\_\_\_\_  
 IV. \_\_\_\_\_  
 V. Revised Scoring \_\_\_\_\_

## HEAD/EYE POSITION

## UPPER FACE

I. Initial Scoring  $(1C \text{ or } 1+4C) + 5 + (6 \text{ or } 7)$  \_\_\_\_\_  
 II. Omission Check  $44C$  \_\_\_\_\_  
 III. Reorganized Scoring  $(1C \text{ or } 1+4C) + 5 + (6 \text{ or } 7) + 44C$  \_\_\_\_\_  
 IV. Co-Occurrence Rule Check (Include 9, 11, 53, 54, 61, 62, 63, 64, 65, 66 if scored) \_\_\_\_\_  
 AUs Numeric Order  $1+4+5+6+7+44$  \_\_\_\_\_  
 Rules:  $5 > 44$   $7 < 44$  \_\_\_\_\_  
 IV.  $(1C \text{ or } 1+4C) + (5 \text{ or } 44C) + (6 \text{ or } 7) + (7 \text{ or } 44C)$  \_\_\_\_\_  
 V. Revised Scoring  $1+4+5+(6 \text{ or } 7)$  \_\_\_\_\_

Special Minimum Requirements  
 With Lower Face  
 4 with 9  
 6 with 9, 10, 12, 13  
 7 with 12, 13  
 41 with 9

FINAL SCORING LOWER FACE: \_\_\_\_\_ HEAD/EYE: \_\_\_\_\_  
 UPPER FACE:  $1+4+5+7$  \_\_\_\_\_ 74. Unscorable (p.10-2)

FINAL FULL FACE SCORE: \_\_\_\_\_  
 73. Entire Head/Face out of view (p. 10-2)

Coder's Name: \_\_\_\_\_ Date: \_\_\_\_\_ Time: \_\_\_\_\_  
 Stimulus: \_\_\_\_\_ Segment: \_\_\_\_\_ Item: \_\_\_\_\_  
 Location: Beginning \_\_\_\_\_ End \_\_\_\_\_

Chapter 3 8

(a)

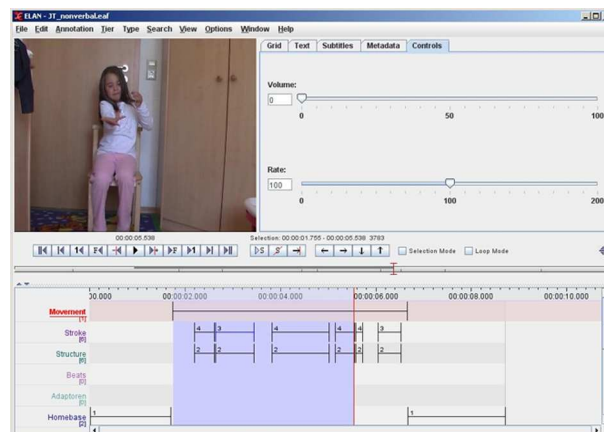


(b)

**Abb. 1.2:** Manuelle Erfassung eines Interaktionsverlaufs: Die Kodierung eines Interaktionsverlaufs erfolgt in Feldbeobachtungen durch Fragebögen (a; Altmann u. a. 2008) teils videogestützt (b; Dörfler u. a. 2001).

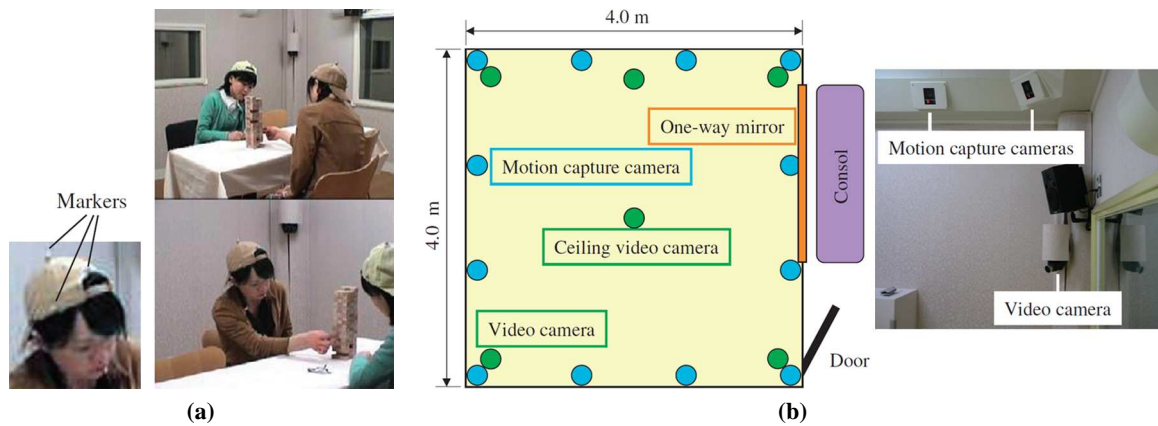


(a)

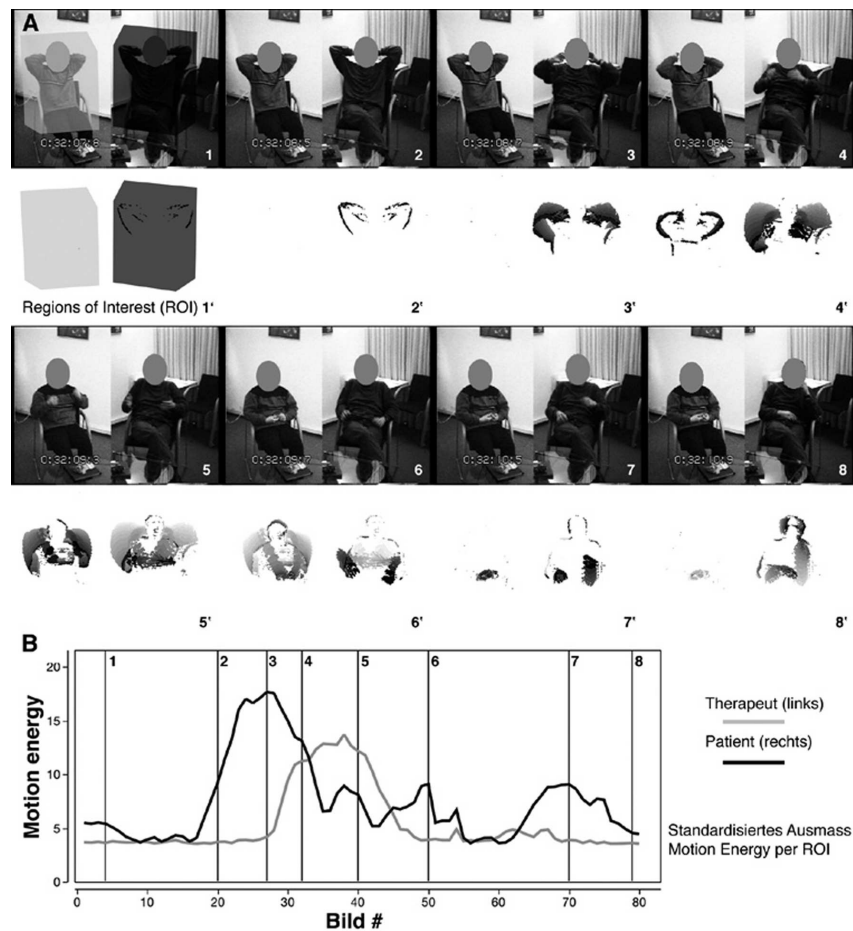


(b)

**Abb. 1.3:** Videoaufnahmen sprachbegleitender Gesten (a; Lausberg & Kryger 2011) lassen sich mit Hilfe der Annotationssoftware ELAN (Max-Planck-Institut für Psycholinguistik in Nijmegen, NL) effektiv analysieren (b; Hogrefe u. a. 2011).



**Abb. 1.4:** Automatisierte Erhebung durch Trackingsysteme (Kawaguchi u. a. 2010): Das typische Bewegungsverhalten das Balancerspiel *Jenga* spielender autistischer Kinder wird über die Verfolgung der mit Zielmarken signalisierten Köpfe (a) durch ein Mehrkamarasystem erfasst (b).



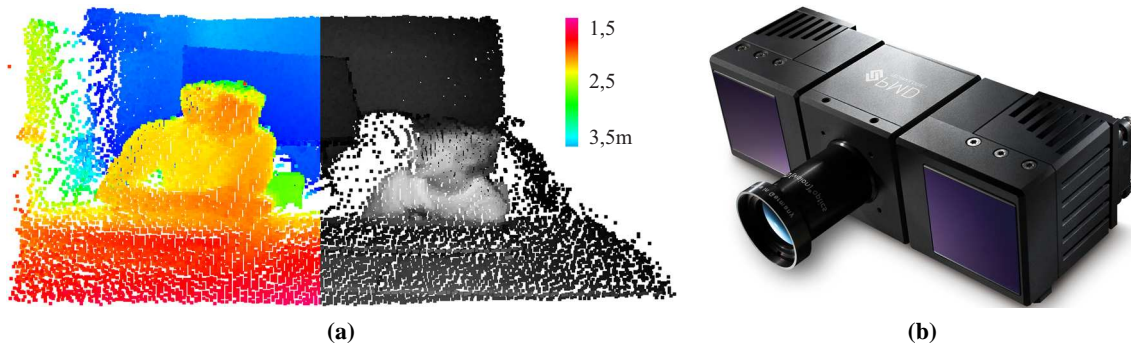
**Abb. 1.5:** Automatisierte Erhebung durch MEA (Ramseyer 2010): Das nonverbale Verhalten wird in dieser Therapiesitzung durch die für definierte Bildbereiche ermittelte Bewegungsenergie quantifiziert. Die Bewegungsenergie wird oben als Differenzbild und unten als Zeitreihe dargestellt.

nalaufwendig. Eine mögliche Alternative ist der Einsatz kommerziell verfügbarer Trackingsysteme, welche spatiotemporal hochaufgelöste Bewegungsvektorfelder liefern, jedoch seitens des Mehraufwands einer multiokularen Experimentalanordnung sowie der Notwendigkeit der Signalisierung für verhaltensanalytische Anwendungen nicht immer geeignet sind. Bei der MEA entfällt die Markierung der Probanden mit auffälligen Sensoren und somit eine Störquelle im Hinblick auf eine möglichst natürliche Interaktion. Prinzipbedingt werden bei diesem Verfahren die Körperbewegungen aber nur indirekt registriert und in Abhängigkeit des betrachteten Bereichs Bewegungen verschiedener Körperteile auf ein skalares Maß reduziert. Neben diesem die Messgenauigkeit empfindlich beeinflussenden Informationsverlust ergeben sich weiterhin geometrische Restriktionen aufgrund des monokularen zentralperspektivischen Aufnahmevorgangs (Abbildungsmaßstab). Die Forderung nach einem flexiblen Verfahren, das spatiotemporal hochauflösende Bewegungsinformationen automatisch generiert und nicht in den Interaktionsverlauf beeinflussend eingreift, bleibt somit weiterhin bestehen.

Hier bietet sich die Photogrammetrie als flexibles Werkzeug für vielfältige Aufgabenstellungen der Bewegungsanalyse an. Die Methoden und Verfahren der Photogrammetrie verstehen sich als ein Teilgebiet der optischen 3D-Messtechnik und 3D-Bildverarbeitung. Sie ermöglichen die zeitaufgelöste, hochgenaue und berührungslose Erfassung und Modellierung beliebiger räumlicher Objekte aus reflektierter oder emittierter Strahlung. Die Nahbereichsphotogrammetrie konnte sich in den letzten Jahrzehnten durch das Aufkommen digitaler Bildaufnahmeverfahren und Bildverarbeitungstechniken sowie die enormen Fortschritte in der Rechnerperformance vielfältige interdisziplinäre Einsatzmöglichkeiten erschließen. Diese konzentrieren sich bislang primär auf Aufgabenstellungen der 3D-Datenakquisition und -analyse für technisch-industrielle Anwendungen oder für Geoinformationssysteme. Die Anwendungsgebiete reichen von Automobil-, Maschinen- und Schiffbau über Architektur, Denkmalpflege und Archäologie bis hin zu Bewegungsanalysen in Bereichen der Strömungsmesstechnik oder Ballistik. Für die Photogrammetrie ergibt sich in der hochaufgelösten quantitativen Erfassung von Bewegungsabläufen in der Analyse nonverbaler Kommunikation in Interaktionsverläufen von Personen ein interessantes neues Anwendungsgebiet.

Die Eignung von Halbleitersensoren für die Akquisition von Bildsequenzen erlaubt die Aufnahme spatiotemporal aufgelöster Daten. Durch automatische Verfahren der Bildverfolgung lassen sich aus Bildsequenzen 2D-Bewegungsvektorfelder berechnen, deren räumliche und zeitliche Auflösung durch die Parameter der aufnehmenden Kamera bestimmt wird. Werden zwei oder mehr synchronisierte Kameras verwendet, ergibt sich die Möglichkeit der 3D-Bewegungsanalyse. Solche Systeme bieten mehrere Anbieter an. Sie werden intensiv in Bereichen wie Biomechanik, Rehabilitation und Filmanimation eingesetzt und bieten entscheidende Vorteile gegenüber konkurrierenden Bewegungsanalyseverfahren wie z. B. Magnetfeldsensornetzen. Bei der Auswertung stereoskopischer Bildsequenzen multiokularer Systeme müssen die Bildpunkte eines interessierenden Körperpunkts sowohl räumlich zwischen zeitgleich aufgenommenen Bildern mehrerer Kameras als auch zeitlich durch die Objektverfolgung einander zugeordnet werden. Die Verfahren der photogrammetrischen 3D-Bewegungsanalyse sind also als spatiotemporale Zuordnungsaufgaben zu charakterisieren, welche

je nach Aufgabenstellung hohe Komplexitäten aufweisen und nicht immer zu fehlerfreien Lösungen führen. In kommerziellen Systemen wird durch die diskrete Signalisierung der Probanden mit bspw. retroreflektierenden Zielmarken dennoch eine zuverlässige und teils echtzeitfähige Arbeitsweise erreicht. Das Messverfahren kann dann aber nicht mehr als berührungslos bezeichnet werden.



**Abb. 1.6:** 3D-Kamerabildsequenzanalyse in der Verhaltensforschung: Am Beispiel der Untersuchung des Konfliktverhaltens zweier Probanden während einer gemeinsamen Aufgabenbewältigung wird in (a) eine Epoche des Interaktionsverlaufs als 3D-Punktwolke mit einer jeweils halbseitigen Überlagerung der Grauwert- und farbkodierten Tiefenwertinformationen dargestellt. Die Daten wurden mit einer 3D-Kamera PMD[vision]<sup>®</sup> CamCube 3.0 akquiriert (b; PMDTec 2010).

3D-Tiefenbildkameras stellen eine interessante Alternative der Datenakquisition dar (Abbildung 1.6). Auf der Basis von Photomischdetektoren nehmen sie zusätzlich zu einem Grauwertbild synchron ein Distanzbild auf, welches für jedes Bildelement eine Entfernungsinformation enthält. Auf diese Weise werden räumlich aufgelöste Oberflächendaten ohne die Notwendigkeit einer stereoskopischen Zuordnung gewonnen. In der 3D-Bewegungsanalyse ergeben sich durch den Wegfall des räumlichen Zuordnungsschritts bezüglich der Komplexität und des Rechenaufwands erhebliche Erleichterungen. 3D-Kameras weisen mit einem Sensorformat von derzeit bis zu  $204^2$  px zwar ein noch vergleichsweise geringes Sensorformat auf, stellen aber aufgrund ihrer maximalen zeitlichen Auflösung von 50 3D-Bildern pro Sekunde sowie ihrer mit einer Kompaktkamera vergleichbaren Handlichkeit ein leistungsfähiges Werkzeug der Datenakquisition für Aufgabenstellungen der 3D-Bewegungsanalyse dar. Durch den Einsatz einer simultan mit hoher Wiederholungsrate Grauwert- und Entfernungsbilder aufnehmenden 3D-Kamera sowie der zielgerichteten Adaptation und Weiterentwicklung photogrammetrischer Verfahren zur Auswertung von 3D-Kamerabildsequenzen kann in der Verhaltensforschung eine deutliche Erhöhung des Auflösungsgrades und der Effizienz der Auswertungsverfahren erreicht werden.

## Zielsetzung

3D-Kameras ermöglichen die simultane Akquisition von Grauwert- und Entfernungsbildern prinzipiell beliebiger Objektszenen. Für jedes Pixel des Sensors wird zusätzlich zur Bildinformation durch Modulationstechniken ein Entfernungswert ermittelt. Als Ergebnis liegt eine 3D-Darstellung des Objektraums mit einer monochromatischen Texturüberlagerung vor. 3D-Kameras haben das Potential als Einzelsensoren multiokulare Aufnahmesysteme zu ersetzen. Sie bieten sich für eine Reihe von Aufgabenstellungen der 3D-Bewegungsanalyse an, bei denen niedrige Sensorkosten und eine hohe zeitliche Auflösung verlangt werden. Zusammen mit geeigneten Verfahren der Bestimmung von 3D-Bewegungsinformation können sie der Photogrammetrie eine Reihe neuer Anwendungsbereiche erschließen.

Im Rahmen der vorliegenden Dissertation soll eine Methode zur 3D-Bewegungsanalyse aus Bildsequenzen von 3D-Kameras entwickelt werden, welche dichte 3D-Bewegungsvektorfelder von Punkten der Körperoberfläche bestimmen kann. Die Entwicklung geht aus von dem bekannten und in der monoskopischen und stereoskopischen Bildsequenzanalyse häufig eingesetzten spatialen und temporalen Zuordnungsverfahren *Least Squares Matching (LSM)*. Auf dessen Basis soll ein integrierter Bewegungsanalysealgorithmus entwickelt werden, welcher das Potential und die Charakteristik von 3D-Kameras berücksichtigt. Insbesondere setzt sich die Arbeit die Umsetzung der folgenden für die Methode geltenden Eigenschaften als Ziel:

- Das primäre Ziel ist die Entwicklung einer geschlossenen Lösung für die Zuordnung von Oberflächensegmenten auf der Zeitachse in den Grauwert- und Entfernungsbildern einer 3D-Kamera. Dabei verspricht die integrierte und auf die Datenstruktur angepasste Formulierung des funktionalen und stochastischen Zusammenhangs zwischen verschiedenartigen Beobachtungsgruppen eine vollständige Ausnutzung des Informationshaushalts und einen Zugewinn an Genauigkeit und Zuverlässigkeit.
- Damit einhergehend besteht die Notwendigkeit zur Anpassung der Parametrisierung. Der ursprüngliche LSM-Ansatz bestimmt sechs Parameter einer 2D-Affintransformation, nämlich Verschiebungen, Rotationen und Maßstabsanpassungen des zuzuordnenden Bildausschnitts. Insbesondere die Verknüpfung der perspektivbedingten Größenänderung des Bildausschnitts im Grauwertbild mit den gemessenen Translationen in Tiefenrichtung ist naheliegend. Zusätzlich soll der Einfluss weiterer Transformationsparameter, welche die Affintransformation nicht erfasst, abgeschätzt und berücksichtigt werden.
- Die zu entwickelnde Methode ist anhand verschiedener Experimentalkonfigurationen hinsichtlich der Genauigkeits- und Zuverlässigkeitssteigerung zu evaluieren.



Im breiten Gebiet der menschlichen Bewegungsanalyse liegen die Vorteile einer 3D-Kamera vor allem in Anwendungen, die durch einen beschränkten Objektraum und durch die Notwendigkeit einer noninvasiven Sensorik und Auswertemethodik gekennzeichnet sind. Dazu zählt z. B. die Erfassung von Körperbewegungen in der Analyse nonverbaler Kommunikation in Interaktionsverläufen von zwei Personen. Als praktisches Beispiel sollen daher konzipierte Szenarien aus der nonverbalen Verhaltensforschung das Potential der entwickelten Methode zur automatischen und berührungslosen Verfolgung von Objektpunkten in 3D-Kamerabildsequenzen aufzeigen.

Bei herkömmlichen Kameras wird über das Modell der zentralperspektiven Abbildung unter der Verwendung der Orientierungsparameter und Bildkoordinaten der Bezug zwischen Bild- und Objektraum hergestellt. Ein ideales Modell der Zentralprojektion wird auch für die Abbildung von durch eine 3D-Kamera aktiv ausgesandter und vom Objekt reflektierter Infrarotstrahlung auf dem Sensor unterstellt. Wird der Sensor für messtechnische Zwecke verwendet, sind die Abweichungen von diesem Idealmodell folglich durch geeignete Kalibrieransätze zu korrigieren. Zusätzlich messen 3D-Kameras die Strecke zwischen dem Projektionszentrum und jedem im Objektraum beleuchteten Oberflächenpunkt. Neben der Modellierung optospezifischer Fehler müssen folglich auch jene auf eine Distanzmessung wirkende Fehlereinflüsse untersucht und berücksichtigt werden.

Für 3D-Kameras als eine relativ junge Messtechnologie existieren noch keine standardisierten Kalibrierverfahren. Einen zweiten Schwerpunkt der vorliegenden Arbeit bildet somit die Sensormodellierung mit insbesondere diesen Zielen:

- Die eine 3D-Kameramessung beeinflussenden Fehlerquellen werden zunächst im Hinblick auf ihre Ursachen und Wirkungen erläutert. Ist der Einfluss hinreichend beschreibbar – und somit auch parametrisierbar – erfolgt die Aufnahme in eine zu entwickelnde Kalibrierroutine.
- Im Vordergrund steht dabei die Entwicklung eines flexiblen Verfahrens, welches ohne die Notwendigkeit des Einsatzes von Messgeräten übergeordneter Genauigkeit unter der Verwendung eines in seiner Dimension handhabbaren und lediglich grob koordinatenmäßig bekannten Referenzfelds alle unbekannt Kalibrierparameter in einem Guss bestimmt.
- Nach dem Vorbild einer photogrammetrischen Selbstkalibrierung sollen alle zur Verfügung stehenden ursprünglichen, also direkt gemessenen, 3D-Kamerainformationen in einem gemeinsamen funktionalen und stochastischen Modell ausgeglichen werden.
- Die entwickelte integrierte Bündelblockausgleichung ist ausführlich zu validieren.

## Gliederung der Arbeit

Die vorliegende Dissertation ist in fünf Hauptteile gegliedert: *Einleitung, Theoretische Grundlagen, Sensormodellierung, Bewegung* sowie *Zusammenfassung und Ausblick*.

**Teil I** zeigt in den **Kapiteln 1 bis 3** die Notwendigkeit der Entwicklung automatischer Verfahren zur Erhebung menschlicher Interaktionsverläufe auf der Basis von **3D-Kameratiefenbildsequenzen** und beschreibt die gesetzten Ziele sowie die Gliederung der Dissertationsarbeit.

**Teil II** bündelt die zum Verständnis der Arbeit notwendigen theoretischen Grundlagen. Im **Kapitel 4** werden ausgewählte Aspekte der geodätischen Ausgleichsrechnung und robusten Parameterschätzung behandelt. **Kapitel 5** beschreibt das geometrische Modell eines Abbildungsvorganges in der Nahbereichsphotogrammetrie, definiert wichtige Koordinatensysteme und erläutert photogrammetrische Mess- und Kalibrierungsprinzipien. Die in der Praxis häufig zur berührungslosen **3D-Objekterfassung** eingesetzten Aufnahmeverfahren werden in **Kapitel 6** kategorisiert und kurz erläutert. Das abschließende **Kapitel 7** stellt Grundlagen aus dem Bereich der Bildsequenzanalyse bereit, definiert die Prinzipien einer automatischen Bildzuordnung und stellt drei wichtige Verfahren vor.

**Teil III** behandelt sämtliche sensorielle Gesichtspunkte im Zusammenhang mit den zum Einsatz gekommenen neuartigen **3D-Tiefenbildkameras**. **Kapitel 8** führt die **3D-Kamera** als flächenhaft distanzmessendes Sensorsystem ein, beschreibt die Systemkomponenten sowie das zugrunde liegende geometrisch-physikalische Messprinzip. Erste Anwendungen von **3D-Kameras** werden ebenso dargelegt wie ein einschätzender Vergleich mit konkurrierenden **3D-Aufnahmeverfahren** der Tiefenbildgenerierung. Im **Kapitel 9** werden die auf eine **3D-Kameramessung** wirkenden Fehler klassifiziert, deren Ursachen aufgedeckt und der Einfluss auf die resultierenden Messergebnisse quantifiziert. Das folgende **Kapitel 10** zeigt den Stand der Forschung von in der Literatur vorgestellten Kalibrieransätzen zur Modellierung solcher Fehler. Es erläutert die verschiedenen Strategien mit ihren prinzipbedingten Vor- und Nachteilen, insbesondere im Hinblick auf das im Rahmen der Arbeit entworfene und im **Kapitel 11** vorgestellte integrierte Kalibrierverfahren.

**Teil IV** diskutiert Methoden und Verfahren zur Erfassung und Analyse von Bewegungen aus Bild-  
daten, präsentiert ein entwickeltes **3D-Kamerabildzuordnungsverfahren** und zeigt dessen Anwendungspotential im Bereich menschlicher Interaktionsanalysen. Konkret werden im **Kapitel 12** interaktive, automatisierte und automatische Verfahren zur Extraktion von Bewegungsinformation aus Bildsequenzen vergleichend dokumentiert und anhand zahlreicher Beispiele aus Wissenschaft und Praxis belegt. Das folgende **Kapitel 13** beschreibt darauf aufbauend den Stand der Forschung auf dem Gebiet der (automatischen) Analyse von **3D-Kamerabildsequenzen**, kategorisiert die seit Aufkommen der jungen Tiefenbildtechnologie publizierten Methoden und Verfahren und bewertet diese

vor allem bezüglich der verwendeten 3D-Kameramessdaten. Ein zur zeitaufgelösten Auswertung von 3D-Kameradaten entwickeltes flächenbasiertes und berührungslos arbeitendes Bildzuordnungsverfahren ist Gegenstand des Kapitels 14. Die funktionalen und stochastischen Zusammenhänge werden besonders unter dem Gesichtspunkt der simultanen Nutzung aller zur Verfügung stehenden Messdaten dokumentiert und anschließend validiert. In Kapitel 15 kommt das integrierte Zuordnungsverfahren an ausgewählten praktischen Szenarien der menschlichen Bewegungs- und Verhaltensanalysen exemplarisch zum Einsatz.

Teil V schließt die Arbeit in den Kapiteln 16 und 17 mit einer Zusammenfassung und einem Ausblick ab.



# II

## Theoretische Grundlagen



## Mathematische Grundlagen

Dieses Kapitel erklärt die für das Verständnis der vorliegenden Arbeit notwendigen mathematischen Grundlagen in den Bereichen ebene und räumliche Koordinatentransformationen (Abschnitte 4.1 bis 4.3), Ausgleichsrechnung (Abschnitt 4.4) und Statistik (Abschnitte 4.5 und 4.6) sowie robuste Schätzverfahren (Abschnitt 4.7).

### 4.1 Ebene Koordinatentransformation

Ein kartesisches zweidimensionales Koordinatensystem  $xy$  kann mit seinen Punkten  $\mathbf{P}(x, y)$  gegenüber einem übergeordneten Zielsystem  $XY$  verdreht, verschoben und skaliert sein (Abbildung 4.1). Durch eine ebene Ähnlichkeitstransformation lassen sich beide Systeme zur Deckung bringen und die Koordinaten der Zielpunkte  $\mathbf{P}(X, Y)$  berechnen.

Die allgemeine ebene Affintransformation berücksichtigt zwei Verschiebungen  $\mathbf{X}_0(X_0, Y_0)$ , zwei unabhängige Winkel  $(\alpha, \beta)$  für die Verdrehung der beiden Koordinatenachsen sowie zwei Maßstabsfaktoren  $(m_X, m_Y)$  in beiden Koordinatenrichtungen. Die 6-Parametertransformation lautet:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} m_X \cdot \cos \alpha & -m_Y \sin(\alpha + \beta) \\ m_X \cdot \sin \alpha & m_Y \cos(\alpha + \beta) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} \quad (4.1)$$

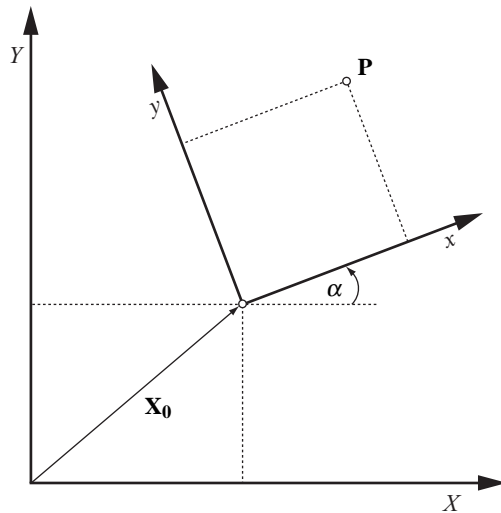
Als eine alternative Schreibweise wird häufig die folgende lineare Form angegeben:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} \quad (4.2)$$

Die Rotationsmatrizen in den Gleichungen (4.1) und (4.2) sind nicht orthogonal; ihre Elemente erfüllen nicht die Orthogonalitätsbedingung. Folglich sind lediglich die Parallelität von Geraden und die Längen von zwei Strecken, die auf parallelen Geraden liegen, invariant gegenüber einer 2D-Affintransformation. Für alle anderen Punkte ändert sich hingegen das Streckenverhältnis. Auch stehen rechtwinklige Geraden nach einer Affintransformation nicht mehr rechtwinklig aufeinander. Ein Quadrat wird bspw. als Parallelogramm abgebildet.

Die allgemein übliche 4-Parameter-Ähnlichkeitstransformation geht aus der ebenen Affintransformation hervor. Auch sie verschiebt das  $xy$ -Ausgangssystem um den Vektor  $\mathbf{X}_0$  in das  $XY$ -Zielsystem. Im Gegensatz zu der oben beschriebenen ebenen Affintransformation verwendet sie jedoch einen einheitlichen Maßstabsfaktor  $m$  und eine orthogonale Rotationsmatrix mit dem Drehwinkel  $\alpha$ . Sie wird in der Gleichung (4.3) für die Transformation vom  $xy$ -Ausgangs- in das übergeordnete  $XY$ -

Zielsystem angeben:



$$\mathbf{X} = m \cdot \mathbf{R} \cdot \mathbf{x} + \mathbf{X}_0$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = m \cdot \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} \quad (4.3)$$

Abb. 4.1: Ebene Ähnlichkeitstransformation.

Die Winkel und Streckenverhältnisse bleiben bei einer ebenen Ähnlichkeitstransformation erhalten. Sie ist somit formtreu und bildet bspw. ein Quadrat als ein Quadrat ab, welches lediglich verschoben, verdreht und skaliert sein kann.

Zur Lösung sind mindestens zwei identische (Pass-)Punkte in beiden Koordinatensystemen notwendig. Liegen mehr gemessene Punkte vor, können die Transformationsparameter statistisch optimal durch Ausgleichung des überbestimmten Gleichungssystems geschätzt werden (Abschnitt 4.4).

## 4.2 Räumliche Koordinatentransformation

Die allgemeine dreidimensionale Affintransformation vollzieht den Übergang von der Ebene in den Raum. Sie bildet in Analogie zur Gleichung (4.2) zwei dreidimensionale kartesische Koordinatensysteme  $xyz$  und  $XYZ$  durch insgesamt 12 Parameter aufeinander ab, nämlich jeweils drei Translationen, Rotationen, Scherungen und Maßstabsfaktoren pro Koordinatenrichtung. Die Transformationseigenschaften können von der ebenen auf die räumliche Affintransformation übertragen werden (Abschnitt 4.1).

Als Spezialfall gilt die räumliche Ähnlichkeitstransformation (auch 3D-Helmerttransformation; Abbildung 4.2). Diese formtreue Abbildung ist die am häufigsten angewendete 3D-Koordinatentransformation (Luhmann 2003). Sie wird durch insgesamt sieben Parameter beschrieben: drei Translationen  $\mathbf{X}_0(X_0, Y_0, Z_0)$ , drei Winkel ( $\omega, \varphi, \kappa$ ) zur Drehung des Systems um jede Achse mittels einer orthogonalen Rotationsmatrix  $\mathbf{R}$  sowie einen für alle Richtungen einheitlichen Maßstabsfaktor  $m$ . Im Vergleich zur 3D-Affintransformation erlaubt dieser reduzierte Parametersatz eine formtreue Abbildung, insbesondere unter der Wahrung der Rechtwinklichkeit der Koordinatenachsen. Die Vorschrift für die Transformation eines Punkts  $\mathbf{P}(x, y, z)$  vom  $xyz$ - ins  $XYZ$ -System lautet:

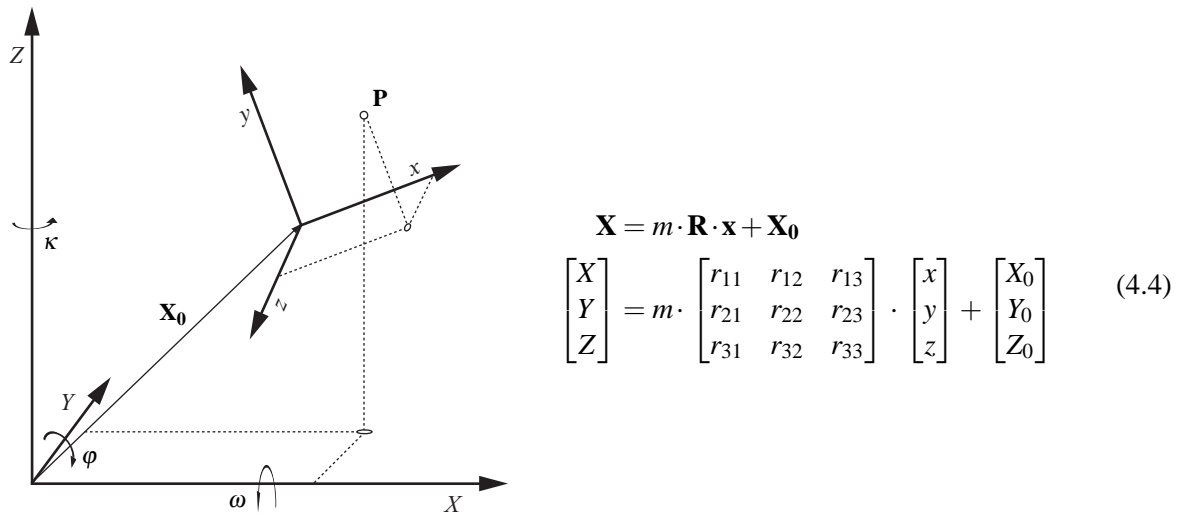


Abb. 4.2: Räumliche Ähnlichkeitstransformation.

Für die eindeutige Berechnung der sieben Transformationsparameter sind sieben Beobachtungen notwendig. Bei z. B. einer Aufnahme in Z-Richtung setzen sich diese aus zwei X- und zwei Y- sowie drei Z-Komponenten von drei räumlich verteilten Passpunkten zusammen. Im Falle einer Überbestimmung können die Beobachtungen bspw. vermittelnd ausgeglichen werden (Abschnitt 4.4).

### 4.3 Räumliche Drehung

Die bereits im vorangegangenen Abschnitt 4.2 eingeführte räumliche Drehung erfolgt im Gegenuhrzeigersinn durch sequentielle Einzeldrehungen um die X-Achse mit dem Winkel  $\omega$ , um die Y-Achse mit dem Winkel  $\varphi$  und um die Z-Achse mit dem Winkel  $\kappa$ . Jede Einzeldrehung wird dabei durch die orthogonalen Rotationsmatrizen  $\mathbf{R}_\omega$ ,  $\mathbf{R}_\varphi$  bzw.  $\mathbf{R}_\kappa$  mathematisch beschrieben:

$$\mathbf{R}_\omega = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{bmatrix} \quad \mathbf{R}_\varphi = \begin{bmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{bmatrix} \quad \mathbf{R}_\kappa = \begin{bmatrix} \cos \kappa & -\sin \kappa & 0 \\ \sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.5)$$

Die in Gleichung (4.4) zur vollständigen Systemrotation definierte räumliche Drehmatrix  $\mathbf{R}$  ergibt sich unter Beachtung der Drehreihenfolge als Produkt der Einzeldrehungen zu:

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_{\omega\varphi\kappa} = \mathbf{R}_\omega \cdot \mathbf{R}_\varphi \cdot \mathbf{R}_\kappa \\ &= \begin{bmatrix} \cos \varphi \cos \kappa & -\cos \varphi \sin \kappa & \sin \varphi \\ \cos \omega \sin \kappa + \sin \omega \sin \varphi \cos \kappa & \cos \omega \cos \kappa - \sin \omega \sin \varphi \sin \kappa & -\sin \omega \cos \varphi \\ \sin \omega \sin \kappa - \cos \omega \sin \varphi \cos \kappa & \sin \omega \cos \kappa + \cos \omega \sin \varphi \sin \kappa & \cos \omega \cos \varphi \end{bmatrix} \end{aligned} \quad (4.6)$$

Die Verwendung von Eulerwinkeln zur Rotationsbeschreibung birgt Nachteile. Vor allem aufgrund

der Mehrdeutigkeiten bei trigonometrischen Funktionen wird in der Praxis häufig auf eine alternative algebraische Darstellungsform mit Quaternionen ausgewichen. Ähnlich der komplexen Zahlen besteht das Zahlensystem der Quaternionen aus einem Real- und einem Imaginärteil. Mit der Hilfe von vier Quaternionen  $q_i$  mit  $i = [1, 4]$  läßt sich eine Drehung im Raum einfach und elegant beschreiben:

$$\mathbf{R} = \begin{bmatrix} q_4^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_3q_4) & 2(q_1q_3 + q_2q_4) \\ 2(q_1q_2 + q_3q_4) & q_4^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_1q_4) \\ 2(q_1q_3 - q_2q_4) & 2(q_2q_3 + q_1q_4) & q_4^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (4.7)$$

mit

$$1 = q_1^2 + q_2^2 + q_3^2 + q_4^2$$

Die geometrische Veranschaulichung von hyperkomplexen Quaternionen ist schwierig, kann aber über eine Umrechnung der Rotationselemente in die übliche trigonometrische Darstellungsform erleichtert werden (Luhmann 2003).

## 4.4 Ausgleichsrechnung

In Naturwissenschaft und Technik werden viele Problemstellungen über die Aufstellung eines geeigneten geometrisch-physikalischen Modells des zu untersuchenden Sachverhalts gelöst. Die unbekannt Parameter dieses Modells sind in der Regel nicht direkt bestimmbar. Anstelle dessen werden Größen beobachtet, also zum Beispiel gemessen, die in einem funktionalen Zusammenhang mit den Unbekannten stehen. Zwei einfache Beispiele im geodätischen Kontext sind die Bestimmung von Höhen aus gemessenen Höhenunterschieden in einem Nivellementnetz oder die lagemäßige Bestimmung eines Neupunkts über Winkel- und Streckenmessungen.

Sobald die Anzahl der Messungen größer als die Anzahl der Modellparameter ist, existiert genau dann eine eindeutige Lösung, wenn alle Beobachtungen fehlerfrei gemessen wurden. In der Praxis ist jedoch jede Beobachtung aufgrund der Unzulänglichkeit der menschlichen Sinne und der Messwerkzeuge fehlerbehaftet. Völlig falsche Messwerte und grobe Fehler werden im Vorfeld durch geeignete Plausibilitätskontrollen eliminiert. Die verbleibenden zufälligen Abweichungen repräsentieren die zu erwartenden Streuungen der Messwerte und bilden die Grundlage für eine Lösung im Sinne der Statistik. Es kommen nun Verfahren der Ausgleichsrechnung zur Anwendung (Abschnitt 4.4.2 und 4.4.3), welche die fehlerbehafteten Beobachtungen ausgleichen und die unbekannt Parameter in funktionalen und stochastischen Modellen schätzen. Das Vorhandensein überschüssiger Beobachtungen ermöglicht darüber hinaus die Berechnung statistischer Qualitätsmaße. Diese sind zur Beurteilung von Genauigkeit und Zuverlässigkeit heranzuziehen (Abschnitt 4.5).

Die folgenden Abschnitte betrachten die grundlegenden Aspekte der Ausgleichsrechnung, welche für das Verständnis der in der Arbeit verwendeten Methoden notwendig sind. Als weiterführende Literatur sind (Reißmann 1976; Niemeier 2002; Koch 2004) zu nennen.

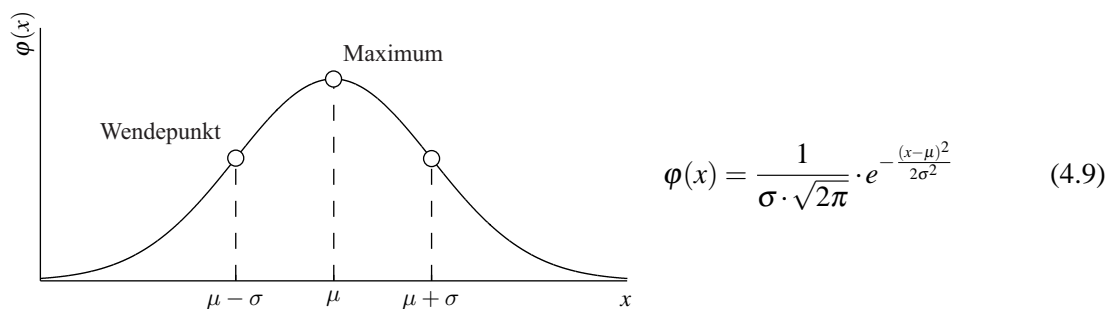
#### 4.4.1 Die Methode der kleinsten Quadrate

Die **Methode der kleinsten Quadrate (MKQ)** oder auch „Schätzung nach L2-Norm“ ist das am häufigsten angewendete Prinzip unter den statistischen Schätzverfahren in den Ingenieurwissenschaften. Nach heutiger Auffassung wurde es ca. 1797 von Carl Friedrich Gauß entwickelt und wenige Jahre später publiziert (**Gauß 1809**).

Eine Ausgleichsaufgabe liegt genau dann vor, wenn die Anzahl  $n$  der Beobachtungen größer ist als die Anzahl  $u$  der zu bestimmenden Parameter:

$$f = n - u \quad (4.8)$$

Redundante Messungen führen somit zu einer Überbestimmung mit  $f > 0$  Freiheitsgraden. Die Aufgabe der **MKQ** ist die Schätzung der wahrscheinlichsten Werte für die gesuchten Parameter unter der Verwendung aller Beobachtungen, wobei deren Verteilungsgesetz bekannt sein muss. In der Praxis wird hier meist eine Normalverteilung der Messwerte  $x$  mit den Varianzen  $\sigma^2$  um ihren Erwartungswert  $\mu$  angenommen. Die Dichtefunktion wird als Gaußsche Glockenkurve  $\varphi(x, \mu, \sigma^2)$  bezeichnet (**Abbildung 4.3**).



**Abb. 4.3:** Kurvenverlauf und Parameter einer Normalverteilung von Zufallsgrößen.

Beobachtungen stellen also stochastische Größen, das heißt Zufallsvariablen dar. Entsprechend dem Prinzip der Methode der kleinsten (Fehler-)Quadrate ist eine Ausgleichung so durchzuführen, dass die Summe der Quadrate ihrer normalverteilten zufälligen Messabweichungen minimiert wird:

$$\Omega = \mathbf{v}^T \mathbf{P} \mathbf{v} \rightarrow \min \quad (4.10)$$

Die zufallsbedingten Widersprüche der Beobachtungen kommen in der Berechnung als Verbesserungen  $\mathbf{v}$  vor. Die Matrix  $\mathbf{P}$  wird als Gewichtsmatrix bezeichnet und im folgenden **Abschnitt 4.4.2** hergeleitet.

Das Ausgleichsmodell beinhaltet ein funktionales und ein stochastisches Modell zur Schätzung der Unbekannten sowie der Verbesserung der Beobachtungen. Das funktionale oder auch geometrische Modell beschreibt den Zusammenhang zwischen den Beobachtungen, Unbekannten und zusätzlich eingeführten Bedingungen untereinander. Das stochastische Modell beschreibt die Vari-

enzen und Kovarianzen der ursprünglichen Beobachtungen, also die Streuung um den wahren Wert entsprechend der Normalverteilung sowie deren Korrelationen untereinander.

#### 4.4.2 Ausgleichung nach vermittelnden Beobachtungen

In den meisten Fällen wird bei der gemeinsamen Bestimmung von mehreren unbekanntem und nicht direkt messbaren Parametern die Ausgleichung nach vermittelnden Beobachtungen angewendet. Sie basiert auf der MKQ (Abschnitt 4.4.1) und wird in ihrem vollständigen Ansatz auch als Gauß-Markov-Modell (GMM) bezeichnet.

##### Funktionales Modell

Gegeben sei folgender funktionaler bzw. geometrischer Zusammenhang zwischen den wahren Messgrößen  $\tilde{\mathbf{L}}$  und den wahren Werten der Parameter  $\tilde{\mathbf{X}}$ :

$$\tilde{\mathbf{L}} = \Phi(\tilde{\mathbf{X}}) = \left[ \Phi(\tilde{X}_1) \quad \Phi(\tilde{X}_2) \quad \dots \quad \Phi(\tilde{X}_n) \right]^T \quad (4.11)$$

Grundsätzlich muss davon ausgegangen werden, dass keine Messung fehlerfrei ist. Für die tatsächlichen (fehlerbehafteten) Beobachtungen  $\mathbf{L}$  gilt Gleichung (4.11) genau dann, wenn ihre Messunsicherheiten  $\sigma_i^2$  durch Verbesserungen  $v_i$  modelliert werden. Ebenso sind die wahren Werte der Parameter nicht bekannt und müssen durch die im Rahmen der Ausgleichung geschätzten Parameterwerte  $\hat{\mathbf{X}}$  ersetzt werden. Es gilt nunmehr für die ausgeglichenen Beobachtungen  $\hat{\mathbf{L}}$  folgendes funktionales Modell:

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v} = \Phi(\hat{\mathbf{X}}) \quad (4.12)$$

Die aufgestellten ursprünglichen Beobachtungs- bzw. Verbesserungsgleichungen (4.12) des funktionalen Modells sind in der Regel nichtlineare Gleichungen. Eine Linearisierung erfolgt durch eine Reihenentwicklung nach Taylor an den Stellen  $(\mathbf{X}^0 + \hat{\mathbf{x}})$ , wobei wegen  $\mathbf{L} \ll \mathbf{v}$  nach dem ersten Glied abgebrochen wird:

$$\begin{aligned} \hat{\mathbf{L}} &= \mathbf{L} + \mathbf{v} = \Phi(\mathbf{X}^0 + \hat{\mathbf{x}}) \\ &= \Phi(\mathbf{X}^0) + \left( \frac{\partial \Phi(\mathbf{X})}{\partial \mathbf{X}} \right)^0 \cdot (\hat{\mathbf{x}} - \mathbf{X}^0) \\ &= \mathbf{L}^0 + \left( \frac{\partial \Phi(\mathbf{X})}{\partial \mathbf{X}} \right)^0 \cdot \hat{\mathbf{x}} \end{aligned} \quad (4.13)$$

mit

$\mathbf{X}^0$  Näherungswerte der Unbekannten

$\hat{\mathbf{x}}$  Korrektur zum Näherungswert  $\mathbf{X}^0$

$\mathbf{L}^0$  Funktionswert, berechnet an der Stelle der Näherungswerte  $\mathbf{X}^0$



Die partiellen Ableitungen der linearisierten Beobachtungsgleichungen (4.13) nach allen Unbekannten bilden die Koeffizienten- oder Designmatrix  $\mathbf{A}$

$$\mathbf{A} = \left( \frac{\partial \Phi(\mathbf{X})}{\partial \mathbf{X}} \right)^0 = \begin{bmatrix} \partial \Phi_1 / \partial \mathbf{X}_1 & \partial \Phi_1 / \partial \mathbf{X}_2 & \cdots & \partial \Phi_1 / \partial \mathbf{X}_u \\ \partial \Phi_2 / \partial \mathbf{X}_1 & \partial \Phi_2 / \partial \mathbf{X}_2 & \cdots & \partial \Phi_2 / \partial \mathbf{X}_u \\ \vdots & \vdots & \ddots & \vdots \\ \partial \Phi_n / \partial \mathbf{X}_1 & \partial \Phi_n / \partial \mathbf{X}_2 & \cdots & \partial \Phi_n / \partial \mathbf{X}_u \end{bmatrix} \quad (4.14)$$

Mit den verkürzten Beobachtungen  $\mathbf{l} = \mathbf{L} - \mathbf{L}^0$  folgt aus (4.13) und (4.14) direkt die linearisierte Form des funktionalen Modells in Matrixdarstellung:

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} = \mathbf{A} \hat{\mathbf{x}} \quad (4.15)$$

#### Stochastisches Modell

Das stochastische Modell enthält die Genauigkeitsinformationen des funktionalen Modells, also die (Mess-)Genauigkeiten der ursprünglichen Beobachtungen  $l_i$  in Form der Standardabweichungen  $\sigma_i$  bzw. der Varianzen  $\sigma_i^2$  sowie mögliche Korrelationen zwischen den Beobachtungen  $l_i$  und  $l_j$ . Diese stochastischen Eigenschaften werden in der Kovarianzmatrix  $\Sigma_{ll}$  der ursprünglichen Beobachtungen zusammengefasst:

$$\Sigma_{ll} = \begin{bmatrix} \sigma_1^2 & \rho_{12} \sigma_1 \sigma_2 & \cdots & \rho_{1n} \sigma_1 \sigma_n \\ \rho_{21} \sigma_2 \sigma_1 & \sigma_2^2 & \cdots & \rho_{2n} \sigma_2 \sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} \sigma_n \sigma_1 & \rho_{n2} \sigma_n \sigma_2 & \cdots & \sigma_n^2 \end{bmatrix} \quad (4.16)$$

Die unterschiedlichen Messgenauigkeiten, und somit auch die Zuverlässigkeit und Wertigkeit einer einzelnen Beobachtung im Ausgleichsprozess, werden über die Gewichtsmatrix  $\mathbf{P}$ , gebildet aus den Gewichten  $p_i$ , berücksichtigt:

$$\mathbf{P} = \mathbf{Q}_{ll}^{-1} = \sigma_0^2 \Sigma_{ll} = \begin{bmatrix} p_{11} & & & \\ & p_{22} & & \\ & & \ddots & \\ & & & p_{nn} \end{bmatrix} = \begin{bmatrix} \sigma_0^2 / \sigma_1^2 & & & \\ & \sigma_0^2 / \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_0^2 / \sigma_n^2 \end{bmatrix} \quad (4.17)$$

mit

$\mathbf{Q}_{ll}$  Kofaktormatrix der ursprünglichen Beobachtungen

Liegen keine a-priori Genauigkeitsinformationen für einzelne Beobachtungen bzw. Beobachtungsgruppen vor, müssen die wahren Varianzen  $\sigma_i^2$  aus Näherungen abgeleitet oder abgeschätzt werden (Abschnitt 4.6). Liegen unkorrelierte Beobachtungen vor (Korrelationskoeffizient  $\rho_{ij} = 0$ ), verein-

facht sich  $\Sigma_{ll}$  zu einer Diagonalmatrix. Bei unkorrelierten und gleichgenauen Beobachtungen ist die Kovarianzmatrix  $\Sigma_{ll}$  der ursprünglichen Beobachtungen eine Einheitsmatrix  $\mathbf{E}$ .

Die in Gleichung (4.17) eingeführte Varianz  $\sigma_0^2$  der Gewichtseinheit ist die konstante und in der Regel unbekannte Varianz einer Beobachtung mit dem Gewicht Eins. Ihre positive Wurzel wird als Standardabweichung  $\sigma_0$  der Gewichtseinheit definiert. Diese theoretische Varianz wird für praktische Ausgleichungsprobleme durch die empirische a-priori Varianz  $s_0^2$  bzw. die Standardabweichung  $s_0$  ersetzt. Im Rahmen der Ausgleichung kann dann ein Schätzwert für die a-posteriori Varianz  $\hat{s}_0^2$  bzw. Standardabweichung  $\hat{s}_0$  bestimmt werden.

### Parameterschätzung

Für die Parameterschätzung im **GMM** wird formal von der linearisierten Form des funktionalen Modells ausgegangen. Die Umstellung der Gleichung (4.15) nach  $\mathbf{v}$  ergibt das Verbesserungsgleichungssystem. Durch ein anschließendes Einsetzen in die Minimumbedingung der Verbesserungsquadratsumme (4.10) gilt die folgende Beziehung:

$$\mathbf{v}^T \mathbf{P} \mathbf{v} = (\mathbf{A} \hat{\mathbf{x}} - \mathbf{l})^T \mathbf{P} (\mathbf{A} \hat{\mathbf{x}} - \mathbf{l}) \rightarrow \min \quad (4.18)$$

Folglich lautet die Lösung dieser Extremwertaufgabe:

$$\frac{\partial \mathbf{v}^T \mathbf{P} \mathbf{v}}{\partial \hat{\mathbf{x}}} = 2 \hat{\mathbf{x}}^T \mathbf{A}^T \mathbf{P} \mathbf{A} - 2 \mathbf{l}^T \mathbf{A} = 0 \quad (4.19)$$

Aufgelöst nach dem Lösungsvektor  $\hat{\mathbf{x}}$  folgt die Berechnungsformel für eine reguläre Ausgleichungsaufgabe im **GMM**:

$$\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{l} = \mathbf{N}^{-1} \mathbf{n} = \mathbf{Q}_{xx} \mathbf{n} \quad (4.20)$$

mit

$\mathbf{N}$  Normalgleichungsmatrix

$\mathbf{n}$  Absolutvektor der Normalgleichungen

$\mathbf{Q}_{xx}$  Kofaktormatrix der Unbekannten

Letztlich können mit den geschätzten Zuschlägen  $\hat{\mathbf{x}}$  die Näherungswerte  $\mathbf{X}^0$  der Unbekannten (iterativ) verbessert werden:

$$\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}} \quad (4.21)$$

#### 4.4.3 Ausgleichung nach vermittelnden Beobachtungen mit Bedingungen zwischen den Unbekannten

Das oben beschriebene **GMM** stellt die Beobachtungsgleichungen (4.12) als Funktion der Parameter dar. Sollen zusätzliche Restriktionen zwischen den Parametern in Form von Bedingungsgleichun-

gen eingeführt werden, wird das **GMM** auf den Allgemeinfeld der Ausgleichsrechnung erweitert und als **Gauß-Helmert-Modell (GHM)** bezeichnet (Lawson & Hanson 1995; Björck 1996; Niemeier 2002; Snow 2002).

Nebenbedingungen können eingeführt werden, wenn a-priori Informationen über die Parameter bzw. Relationen der Parameter zueinander vorliegen. Beispiele hierfür sind ein definierter Abstand zwischen zwei ausgeglichenen Messpunktkoordinaten oder die Forderung, dass einige der ausgeglichenen Messpunkte auf einer Kugel liegen. Durch diese jeweils anwendungsspezifisch aufgestellten Nebenbedingungen wird das Gesamtmodell besser an den tatsächlich vorliegenden Sachverhalt angelehnt und genauer modelliert. Die gesamte Ausgleichung wird stabiler, da weniger gute Beobachtungen durch gute Beobachtungen nebst Restriktionen gestützt werden. Darüber hinaus kann ein möglicher Rangdefekt der Normalgleichungsmatrix  $\mathbf{N}$  durch die Einführung von zusätzlichen Bedingungen korrigiert werden. Ein Rangdefekt entsteht, sobald Beobachtungen zur eindeutigen Bestimmung der Unbekannten fehlen.

Die Aufstellung der Verbesserungsgleichungen entspricht der Gleichung (4.12). Zusätzlich kommt eine Anzahl  $b$  von in der Regel nichtlinearen Bedingungsgleichungen hinzu:

$$\Gamma(\tilde{\mathbf{X}}) = \left[ \Gamma(\tilde{X}_1) \quad \Gamma(\tilde{X}_2) \quad \dots \quad \Gamma(\tilde{X}_r) \right]^T \quad (4.22)$$

Die zusätzlichen Bedingungsgleichungen erhöhen den Redundanzanteil des Gleichungssystems:

$$f = n - u + b \quad (4.23)$$

Die Linearisierung der Restriktionen an der Stelle der Näherungswerte  $\mathbf{X}^0$  führt zur Koeffizientenmatrix  $\mathbf{B}$  der Bedingungsgleichungen:

$$\mathbf{B} = \left( \frac{\partial \Gamma(\mathbf{X})}{\partial \mathbf{X}} \right)^0 \quad (4.24)$$

Das linearisierte funktionale Modell für den Allgemeinfeld der Ausgleichsrechnung (**GHM**) lautet somit:

$$\begin{aligned} \mathbf{l} + \mathbf{v} &= \mathbf{A}\hat{\mathbf{x}} \\ \mathbf{w} &= \mathbf{B}\hat{\mathbf{x}} \end{aligned} \quad (4.25)$$

Ziel ist nun die Bestimmung der Unbekannten unter Minimierung der Verbesserungsquadratsumme  $\Omega$  und Einhaltung der aufgestellten Restriktionsgleichungen. Die Minimumbedingung (4.10) des **GMM** erweitert sich zu:

$$\mathbf{v}^T \mathbf{P} \mathbf{v} + 2\mathbf{k}(\mathbf{B}\hat{\mathbf{x}} + \mathbf{w}) \rightarrow \min \quad (4.26)$$

Der Widerspruchsvektor  $\mathbf{w}$  enthält die Abweichungen der Funktionswerte der Bedingungsgleichun-

gen zu Null. Die numerischen Werte der Korrelaten  $\mathbf{k}$  (Lagrange-Multiplikatoren) sind meist nicht von Interesse.

Das erweiterte Normalgleichungssystem mit der Gewichtsmatrix  $\mathbf{P}_B$  der Restriktionen vereint die aufgestellten Beobachtungsgleichungen  $\Phi$  sowie die Bedingungsgleichungen  $\Gamma$  in einem Gleichungssystem und führt zur Lösung der Ausgleichungsaufgabe:

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T \mathbf{P} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & -\mathbf{P}_B^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \mathbf{P} \mathbf{l} \\ \mathbf{w} \end{bmatrix} \quad (4.27)$$

Alternativ können (geometrische) Bedingungen auch als fiktive Beobachtungen  $\mathbf{l}_B$  dem funktionalen Modell (4.15) angefügt werden:

$$\hat{\mathbf{l}} = \begin{bmatrix} \mathbf{l} \\ \mathbf{l}_B \end{bmatrix} + \begin{bmatrix} \mathbf{v} \\ \mathbf{v}_B \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}_B \end{bmatrix} \hat{\mathbf{x}}_B \quad (4.28)$$

Dieser nicht strenge Ansatz ist in der heutigen Praxis vor allem aufgrund seiner enormen Flexibilität in der Gewichtung weit verbreitet (Niemeier 2002). Es kann bspw. unterschieden werden, ob es sich um strenge oder nur genähert einzuhaltende Restriktionen handelt.

## 4.5 Statistische Qualitätsmaße

### 4.5.1 Beobachtungsfehler

Keine Messung ist fehlerfrei, da sie stets durch Mess- bzw. Beobachtungsfehler beeinflusst ist. Prinzipiell wird zwischen drei Arten von Beobachtungsfehlern unterschieden (Kahmen 1997):

- Grobe Fehler im Datenmaterial entstehen meist durch mangelnde Sorgfalt während des Messprozesses und sind durch Plausibilitätskontrollen und Mehrfachmessungen im Vorfeld der Ausgleichung auszuschließen. Ein nachträgliches Aufdecken ist meist schwierig, teilweise aber durch die Anwendung robuster Verfahren noch möglich (Abschnitt 4.7).
- Systematische Fehler werden auch als Modellfehler bezeichnet. Sie entstehen, wenn die physikalischen oder geometrischen Beziehungen zwischen den Beobachtungen und den Unbekannten fehlerhaft sind, also das aufgestellte funktionale Modell nicht der Realität entspricht. Die Messergebnisse müssen vor der Ausgleichung frei von systematischen Fehlern sein (Abschnitt 4.5.5).
- Zufällige Fehler sind die Summe der nach Elimination von groben und systematischen Fehlern verbleibenden unbeherrschbaren Elementarfehler wie z. B. die unkontrollierbaren Änderungen der äußeren Einflüsse. Sie repräsentieren die zu erwartende Streuung der Messwerte um den Erwartungswert mit änderndem Betrag und Vorzeichen. Ihre zufällige Verteilung entspricht der Normalverteilung (Abschnitt 4.4.1) und bildet die Grundlage für die Anwendung

statistischer Ausgleichungsverfahren und die Ableitung von a-posteriori Genauigkeitsmaßen (Abschnitt 4.5.2).

#### 4.5.2 Genauigkeitsmaße

Die empirisch ermittelte a-posteriori Standardabweichung der Gewichtseinheit kann aus allen Verbesserungen  $\mathbf{v}$  angegeben werden:

$$\hat{s}_0 = \sqrt{\frac{\mathbf{v}^T \mathbf{P} \mathbf{v}}{f}} \quad (4.29)$$

Alle weiteren Genauigkeitsmaße können nun aus der a-posteriori Varianz  $\hat{s}_0^2$  der Gewichtseinheit und der Kofaktormatrix  $\mathbf{Q}_{xx}$  bzw. der Kovarianzmatrix  $\mathbf{\Sigma}_{xx}$  der Unbekannten abgeleitet werden. Die Matrizen ergeben sich laut Gleichung (4.20) zu:

$$\mathbf{\Sigma}_{xx} = \hat{s}_0^2 \mathbf{Q}_{xx} = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1u} \\ q_{21} & q_{22} & \cdots & q_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nu} \end{bmatrix} = \begin{bmatrix} \hat{s}_1^2 & \rho_{12} \hat{s}_1 \hat{s}_2 & \cdots & \rho_{1n} \hat{s}_1 \hat{s}_n \\ \rho_{21} \hat{s}_2 \hat{s}_1 & \hat{s}_2^2 & \cdots & \rho_{2n} \hat{s}_2 \hat{s}_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} \hat{s}_n \hat{s}_1 & \rho_{n2} \hat{s}_n \hat{s}_2 & \cdots & \hat{s}_n^2 \end{bmatrix} \quad (4.30)$$

Die a-posteriori Standardabweichung der ursprünglichen Beobachtungen ergibt sich über eine Skalierung der a-priori vorgegebenen Standardabweichungen  $s_i$  der ursprünglichen Beobachtungen mit der a-posteriori Standardabweichung  $\hat{s}_0$  der Gewichtseinheit:

$$\hat{s}_{i_i} = \hat{s}_0 \sqrt{q_{i_i i_i}} \quad (4.31)$$

Die Kofaktormatrix  $\mathbf{Q}_{\hat{I}\hat{I}}$  der ausgeglichenen Beobachtungen ermöglicht die Angabe der Standardabweichung der ausgeglichenen Beobachtungen  $\hat{l}_i$ :

$$\hat{s}_{\hat{l}_i} = \hat{s}_0 \sqrt{q_{\hat{l}_i \hat{l}_i}} \quad (4.32)$$

mit

$$\mathbf{Q}_{\hat{I}\hat{I}} = \mathbf{A} \mathbf{Q}_{xx} \mathbf{A}^T$$

Die a-posteriori Standardabweichung der einzelnen geschätzten Unbekannten  $\hat{x}_i$  ist gegeben durch:

$$\hat{s}_{\hat{x}_i} = \hat{s}_0 \sqrt{q_{x_i x_i}} \quad (4.33)$$

Die a-posteriori Standardabweichung einer Funktion  $F_i$  der geschätzten Unbekannten  $\hat{x}_i$  ist definiert als:

$$\hat{s}_{F_i} = \hat{s}_0 \sqrt{q_{F_i F_i}} \quad (4.34)$$

mit

$$\mathbf{Q}_{FF} = \mathbf{f}\mathbf{Q}_{xx}\mathbf{f}^T, \quad \mathbf{f} = \left(\frac{\partial F(\mathbf{X})}{\partial \mathbf{X}}\right)^0$$

### 4.5.3 Zuverlässigkeitsmaße

Objektive Kriterien zur Beurteilung der Zuverlässigkeit stammen vor allem aus [Baardas](#) Zuverlässigkeitstheorie (1968). Sie können aus der Kofaktormatrix  $\mathbf{Q}_{vv}$  der Verbesserungen abgeleitet werden und bieten Möglichkeiten zur Kontrolle der Beobachtungen sowie deren Auswirkungen auf die Parameter:

$$\mathbf{Q}_{vv} = \mathbf{Q}_{ll} - \mathbf{A}\mathbf{Q}_{xx}\mathbf{A}^T \quad (4.35)$$

Neben der Aufnahmeconfiguration, also der geometrischen Messanordnung, hängt die Zuverlässigkeit von der Anzahl der überschüssigen Beobachtungen ab. Entsprechend der Gleichung (4.8) setzt sich die Gesamtredundanz einer Ausgleichung aus den Redundanzanteilen  $r_i$  einer Beobachtung  $l_i$  zusammen:

$$f = n - u = \text{spur}(\mathbf{Q}_{vv}\mathbf{P}) = \text{spur}(\mathbf{R}) = \sum r_i \quad (4.36)$$

Der Redundanzanteil  $r_i$  ( $0 \leq r_i \leq 1$ ) gibt an, welcher Anteil der Beobachtung zur Berechnung der Parameter verwendet wird und welcher Anteil überschüssig ist. Ein  $r_i$  deutlich größer als der mittlere Redundanzanteil  $\bar{r}_i = 1 - u/n$  weist auf eine gute Kontrollierbarkeit der Beobachtung hin. Eine Beobachtung wird hingegen nicht bzw. nur schlecht von anderen Beobachtungen kontrolliert, wenn ihr  $r_i$  wesentlich kleiner als  $\bar{r}_i$  ist. Ein grober Fehler ist in diesem Fall nicht aufdeckbar.

Ein weiteres Zuverlässigkeitsmaß ist die oft als Testgröße zur Aufdeckung grober Fehler herangezogene normierte Verbesserung  $w_i$  einer Beobachtung  $l_i$  (vgl. [Abschnitt 4.5.4](#)):

$$w_i = \frac{v_i}{\hat{s}_{v_i}} \quad (4.37)$$

mit

$$\hat{s}_{v_i} = \hat{s}_0 \sqrt{q_{v_i v_i}} = \hat{s}_{l_i} \sqrt{r_i} \quad \text{Standardabweichung einer Verbesserung}$$

### 4.5.4 Ausreißerdetektion

Grobe Fehler im Beobachtungsmaterial werden auch als Ausreißer bezeichnet. Sie können bei praktisch jeder Messanordnung auftreten und stehen in keinem Zusammenhang mit der Messgenauigkeit. Die [MKQ](#) hat die Eigenschaft, den Einfluss von Messabweichungen optimal auf alle Beobachtungen zu verteilen. Dementsprechend bewirken grobe Fehler eine Änderung der angenommenen Normalverteilung der Verbesserungen und beeinflussen das Ergebnis der Ausgleichungsaufgabe negativ. Sie sind daher vor der Ausgleichung zu beseitigen ([Abschnitt 4.5.1](#)).

Eine einfache Art der Ausreißeridentifikation ist die Analyse der verkürzten Beobachtungen **I** ([Abschnitt 4.4.2](#)). Liegen a-priori gute Näherungen für die Unbekannten vor, kann die Differenz zwischen den Beobachtungen und den Funktionswerten, berechnet an der Stelle der Näherungswerte, bereits Aufschluss über die Qualität des einzelnen Messwerts geben.

Iterative Verfahren untersuchen nach jedem Durchgang einer Ausgleichung das Datenmaterial sukzessive nach grob fehlerhaften Beobachtungen. Diese Suche wird nach ([Baarda 1968](#)) auch als *Data Snooping* bezeichnet und basiert auf der Annahme, dass sich nur ein grober Fehler in den Beobachtungen befindet. Die in [Abschnitt 4.5.3](#) berechneten normierten Verbesserungen  $w_i$  werden mit einem **Grenzwert für das Data Snooping (GWDS)** verglichen. Ist  $w_i$  größer als der Grenzwert, liegt vermutlich ein grober Fehler vor. In der Praxis ist ab einem **GWDS** von 2,5 ein grober Fehler möglich; ab einem **GWDS** größer 4,0 ist ein grober Fehler sehr wahrscheinlich ([Möser u. a. 2000](#)). Da grobe Fehler die Verbesserungen von eigentlich fehlerfreien Beobachtungen negativ beeinflussen können, ist immer nur die Beobachtung mit der größten  $w_i$  zu eliminieren. Eine detaillierte Übersicht weiterer statistischer Testverfahren findet sich in ([Jäger u. a. 2005](#)).

Die vorgestellten Verfahren zur Ausreißerdetektion setzen voraus, dass sich wenige bzw. nur ein grober Fehler im Beobachtungsmaterial befinden. Sie versagen, sobald eine Messreihe mit vielen Ausreißern behaftet ist. In solch einem Fall können Verfahren der robusten Parameterschätzung vor der eigentlichen Ausgleichung zur Anwendung kommen ([Abschnitt 4.7](#)).

#### 4.5.5 Test einer Normalverteilung

Nachdem grobe Fehler im Datenmaterial detektiert und eliminiert wurden, ist zu untersuchen, ob die Verbesserungen der verbleibenden Beobachtungen normalverteilt sind. Dies ist dann der Fall, wenn keine zusätzlichen systematischen Fehler in den Messreihen auftreten. Zur Prüfung einer Normalverteilung kann neben einer visuellen Kontrolle des Histogramms der Verbesserungen ein statistischer Test herangezogen werden.

Es werden bspw. mit einem  $\chi^2$ -Anpassungstest die Verteilungseigenschaften einer Grundgesamtheit angegeben und untersucht, ob die Verteilungsfunktion einer Normalverteilung  $\varphi(x, \mu, \sigma^2)$  entspricht (Nullhypothese  $H_0$ ; [Abschnitt 4.4.1](#)). Die Prüfgröße  $\hat{\chi}^2$  wird berechnet aus der Summe der normierten Differenzen der Häufigkeiten der empirischen und theoretischen Verteilung, geteilt durch die zu erwartenden Häufigkeiten ([Reißmann 1976](#)). Ist die Prüfgröße  $\hat{\chi}^2$  kleiner als ein gewähltes Quantil der  $\chi^2$ -Verteilung, wird die Nullhypothese angenommen. Die Verteilung der Verbesserungen entspricht dann einer Normalverteilung; systematische Fehler können nicht nachgewiesen werden.

#### 4.5.6 Zusätzliche Parameter

##### Korrelationen

Die Korrelationen zwischen den geschätzten Parametern können ebenfalls als Maß zur Bewertung der Qualität des Ausgleichungsergebnisses herangezogen werden. Die Kovarianzmatrix  $\Sigma_{xx}$  der Un-

bekannten beinhaltet die zur Analyse notwendigen Korrelationskoeffizienten  $\rho_{ij}$ . Es gilt somit:

$$\rho_{ij} = \frac{\hat{s}_{\hat{x}_i \hat{x}_j}}{\hat{s}_{\hat{x}_i} \hat{s}_{\hat{x}_j}} \quad -1 \leq \rho_{ij} \leq +1 \quad (4.38)$$

Lineare Abhängigkeiten zwischen zwei Unbekannten liegen vor, wenn der Korrelationskoeffizient nahe Eins liegt. Das funktionale Modell sollte in diesem Fall geprüft und ggf. ein Parameter eliminiert werden, um Singularitäten bei der Lösung des Normalgleichungssystems zu vermeiden.

### Bestimmbarkeit

Die Hauptdiagonalelemente  $d_{ii}$  der Matrix  $\mathbf{D}$  sind ein Maß für die Bestimmbarkeit eines Parameters  $x_i$ :

$$\mathbf{D} = \mathbf{E} - [\text{diag}(\mathbf{N}) \cdot \text{diag}(\mathbf{N}^{-1})]^{-1} \quad (4.39)$$

mit

$\mathbf{E}$  Einheitsmatrix

Liegt die Bestimmbarkeit  $b_{ii}$  bei einem Wert größer 0,85 (Kraus 1996), verschlechtert sich die Konditionierung des Normalgleichungssystems. Der Parameter  $x_i$  ist weniger gut bestimmbar und sollte aus dem funktionalen Modell ausgeschlossen werden.

### Test auf Signifikanz

Die Signifikanz eines geschätzten Parameters  $\hat{x}_i$  kann durch einen Hypothesentest überprüft werden (Kraus 1996). In der Nullhypothese  $H_0$  wird formuliert, dass ein Parameter nicht signifikant und somit seine Berücksichtigung im funktionalen Modell nicht sinnvoll ist. Zur Berechnung der Testgröße wird  $\hat{x}_i$  ins Verhältnis zu seiner a-posteriori Standardabweichung  $\hat{s}_{\hat{x}_i}$  gesetzt:

$$t_{x_i} = \frac{\hat{x}_i}{\hat{s}_{\hat{x}_i}} \quad (4.40)$$

Ist der Quotient kleiner als ein gewähltes Quantil der t- bzw. Student-Verteilung, wird die Nullhypothese angenommen und der Parameter beim nächsten Durchlauf der Ausgleichung eliminiert.

## 4.6 Varianzkomponentenschätzung

### 4.6.1 Motivation

Das stochastische Modell einer Ausgleichsrechnung wird durch die Kovarianzmatrix  $\mathbf{\Sigma}_{ll}$  der Beobachtungen aus der Gleichung (4.16) beschrieben und beinhaltet die unterschiedlichen Genauigkei-



ten der Messungen in Form von Gewichten:

$$p_i = \frac{s_0^2}{s_i^2} \quad (4.41)$$

Die Gewichtsmatrix  $\mathbf{P}$  aus der Gleichung (4.17) kann als eine Einheitsmatrix  $\mathbf{E}$  definiert werden, wenn gleichartige, also gleichgenaue und unkorrelierte Beobachtungen vorliegen. In diesem Fall erhält jede Beobachtung das Gewicht  $p_i := 1$  und wird hinsichtlich ihrer Zuverlässigkeit und Wertigkeit im Ausgleichsprozess gleich stark berücksichtigt. Bei einer Kombination von unterschiedlichen Datentypen in einer integrierten Ausgleichung ist dieser Ansatz wenig sinnvoll. Zum einen wird nicht der gesamte Informationshaushalt des heterogenen Datenmaterials ausgeschöpft, zum anderen kann eine falsche Gewichtsabstimmung zu einem schlechten Konvergenzverhalten des Parametervektors führen. In den bisherigen Ausführungen der Abschnitte 4.4 und 4.5 wurde davon ausgegangen, dass die a-priori Varianzen  $s_i^2$  der Beobachtungen  $l_i$  und somit auch ihre Gewichte  $p_i$  bekannt sind oder aus Näherungswerten abgeleitet werden. Liegen jedoch keine adäquaten Genauigkeitsinformationen bzw. Informationen über die Genauigkeitsverhältnisse verschiedener Beobachtungsgruppen zueinander vor, ist eine Bestimmung der Varianzen im Zuge einer **Varianzkomponentenschätzung (VKS)** möglich.

Die **VKS** erweitert das stochastische Modell in der Art, dass gleichgenaue Beobachtungen zu einer Beobachtungsgruppe  $i$  mit einer eigenen (unbekannten) Varianzkomponente  $s_i^2$  zusammengefasst werden<sup>1</sup>. Die Kovarianzmatrix  $\Sigma_{ll}$  der Beobachtungen kann über diesen Gruppenvarianzansatz in additive Komponenten  $\Sigma_{ii}$  zerlegt werden, die bis auf die zu schätzenden Varianzen von der Struktur her bekannt sind. Eine weitere Aufspaltung der einzelnen Kovarianzmatrizen zur Berücksichtigung von verschiedenen Fehleranteilen innerhalb der Gruppe führt zur Schätzung von adaptiven Varianzkomponenten. Dies kann bspw. bei der Modellierung eines konstanten und eines linear von der Distanz abhängigen Fehleranteils bei der Streckenmessung der Fall sein. Die Schätzung von Kovarianzen ist prinzipiell ebenfalls möglich. In der Praxis ist dies aber nicht sinnvoll, da sowohl a-priori die Struktur der Kovarianzen bekannt und darüber hinaus die Redundanz der Ausgleichungsaufgabe sehr hoch sein muss (Sieg & Hirsch 2000a). In den folgenden Abschnitten werden somit unkorrelierte Beobachtungen angenommen.

Das theoretische Konzept der **VKS** wurde 1967 erstmals von Kubik vorgestellt und später von Förstner (1979) und Oswald (1992) optimiert. Der praktische Berechnungsablauf ist in (Niemeier 2002; Koch 2004) dokumentiert. Im Kontext einer Anwendung in ingenieurgeodätischen Netzen wird das Verfahren auch sehr ausführlich in (Sieg & Hirsch 2000a, 2000b) beschrieben. Weitere praktische Implementierungen werden bspw. bei der Feldprüfung geodätischer Instrumente (Klein 2001) oder bei der gemeinsamen Auswertung terrestrischer Laserscannerdaten und photogrammetrischen Bilddaten (Schneider 2009) realisiert. Auch im Rahmen der vorliegenden Arbeit kam die **VKS** zur optimalen Aufstellung stochastischer Modelle heterogener Messdaten zum Einsatz (Kapitel 11

<sup>1</sup> Ab dieser Stelle des Abschnitts 4.6 bezeichnet der Index  $i$  eine Gruppe gleichartiger Beobachtungen.

und 14).

#### 4.6.2 Modell der Varianzkomponentenschätzung

Zur einfachen Schätzung von Varianzkomponenten werden die Varianzen gleichartiger Beobachtungen zu einer Gruppenvarianz  $s_i^2$  der Beobachtungsart  $i$  zusammengefasst. Existieren insgesamt  $g$  Gruppen, kann die Kovarianzmatrix  $\Sigma_{ll}$  der Beobachtungen in einzelne Untermatrizen  $\Sigma_{ii}$  zerlegt werden:

$$\Sigma_{ll} = \begin{bmatrix} s_1^2 \mathbf{E} & & & \\ & s_2^2 \mathbf{E} & & \\ & & \ddots & \\ & & & s_g^2 \mathbf{E} \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & & & \\ & \Sigma_{22} & & \\ & & \ddots & \\ & & & \Sigma_{gg} \end{bmatrix} \quad (4.42)$$

Die Formulierung aus Gleichung (4.42) ermöglicht eine weitere Zerlegung der Kovarianzmatrix  $\Sigma_{ll}$  der Beobachtungen in eine additive Komponente  $\mathbf{V}_i$  pro Beobachtungsart:

$$\Sigma_{ll} = \sum_{i=1}^g \mathbf{V}_i = \sum_{i=1}^g \begin{bmatrix} \Sigma_{ii} & \\ & \mathbf{0} \end{bmatrix} \quad (4.43)$$

Die VKS ist ein iteratives Verfahren. Für eine erste Iteration  $n = 1$  wird die Gleichung (4.43) mit Näherungswerten der zu schätzenden Varianzkomponenten  $s_i^2$  initialisiert. Diese stammen aus a-priori Wissen in Form von z. B. Erfahrungswerten oder aus Herstellerangaben. Im Rahmen des in Abschnitt 4.6.3 detailliert dokumentierten Berechnungsschemas lassen sich nun die Faktoren  $(\hat{\alpha}_i^2)_n$  für jeden weiteren Durchlauf  $n$  abschätzen. Die Varianzkomponenten werden sukzessive durch Multiplikation mit den Faktoren verbessert:

$$\hat{s}_i^2 = s_i^2 \prod_{n=1}^m (\hat{\alpha}_i^2)_n \quad (4.44)$$

Dieser Vorgang wird so lange wiederholt, bis nach  $m$  Iterationen alle Faktoren  $\hat{\alpha}_i^2$  nahe Eins liegen, also keine signifikanten Änderungen der abgeschätzten Varianzkomponenten  $\hat{s}_i^2$  zu erwarten sind.

Zusätzlich werden die Varianzen  $V\{\hat{s}_i^2\}$  für jede  $\hat{s}_i^2$  angegeben, um die Qualität des geschätzten stochastischen Modells beurteilen zu können.

#### 4.6.3 Berechnungsablauf

Für jede Iteration der VKS wird die symmetrische Matrix  $\mathbf{W}$  eingeführt:

$$\mathbf{W} = \Sigma_{ll}^{-1} - \Sigma_{ll}^{-1} \mathbf{A} (\mathbf{A}^T \Sigma_{ll}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \Sigma_{ll}^{-1} \quad (4.45)$$

Weiterhin werden der Vektor  $\mathbf{q}$

$$\mathbf{q} = [\mathbf{1}^T \mathbf{WV}_i \mathbf{W}\mathbf{1}] \quad \text{mit} \quad i = [1, g] \quad (4.46)$$

und die ebenfalls symmetrische Matrix  $\mathbf{S}$

$$\mathbf{S} = [\text{sp}(\mathbf{WV}_i \mathbf{WV}_j)] \quad \text{mit} \quad i, j = [1, g] \quad (4.47)$$

berechnet, um mit ihnen die Verbesserungsfaktoren  $\hat{\alpha}_i^2$  zu schätzen:

$$\hat{\boldsymbol{\alpha}} = [\hat{\alpha}_1^2, \hat{\alpha}_2^2, \dots, \hat{\alpha}_g^2]^T = \mathbf{S}^{-1} \mathbf{q} \quad (4.48)$$

Entsprechend der Gleichung (4.44) sind die Varianzkomponenten  $s_i^2$  bzw. die additiven Komponenten  $\mathbf{V}_i$  zu verbessern und als Näherungen für die nächste Iteration ( $n+1$ ) aufzustellen:

$$\begin{aligned} (s_i^2)_{n+1} &= (\hat{\alpha}_i^2)_n \cdot (s_i^2)_n \\ (\mathbf{V}_i)_{n+1} &= (\hat{\alpha}_i^2)_n \cdot (\mathbf{V}_i)_n \end{aligned} \quad (4.49)$$

Es ist nun so lange zu iterieren, bis nach der  $m$ -ten Iteration für alle Verbesserungsfaktoren  $(\hat{\alpha}_i^2)_m = 1 \pm \varepsilon$  gilt. Die Änderungen der geschätzten Varianzkomponenten  $\hat{s}_i^2$  liegen dann unter einem vorgegebenen Schwellenwert  $\varepsilon$  und werden als nicht signifikant erachtet.

Für die a-priori Varianz  $s_0^2$  der Gewichtseinheit kann eine beliebige Konstante angesetzt werden. Hierbei ist lediglich zu beachten, dass die Gewichte der einzelnen Beobachtungsgruppen möglichst nahe Eins liegen, um die numerische Stabilität der Gewichtsmatrix zu gewährleisten (Gleichung 4.41). Ein anschließender Vergleich der a-priori Varianz  $s_0^2$  der Gewichtseinheit mit der nach der Gleichung (4.29) berechneten a-posteriori Varianz  $\hat{s}_0^2$  der Gewichtseinheit zeigt an, ob die geschätzte Gewichtung zu optimistisch ( $\hat{s}_0^2 > s_0^2$ ) oder zu pessimistisch ( $\hat{s}_0^2 < s_0^2$ ) ausfällt. Ist  $\hat{s}_0^2 = s_0^2$ , fällt also die a-priori Standardabweichung  $s_l$  der ursprünglichen Beobachtungen mit der a-posteriori Standardabweichung  $\hat{s}_l$  der ursprünglichen Beobachtungen zusammen, liegen optimal veranschlagte Gewichte vor (vgl. Gleichung 4.31).

Abschließend ist nach (Koch 2004) die Berechnung der Varianzen  $V\{\hat{s}_i^2\}$  der Varianzkomponenten  $\hat{s}_i^2$  möglich:

$$\hat{V}\{\hat{s}_i^2\} = V\{(\hat{\alpha}_i^2)_m (s_i^2)_m\} = (s_i^2)_m^2 V\{(\hat{\alpha}_i^2)_m\} \quad (4.50)$$

mit

$$V\{\hat{\alpha}_i^2\} = 2s_{l_i l_i} \quad \text{und} \quad \mathbf{S}^{-1} = (s_{l_i l_j})$$

Aspekte einer effektiven numerischen Berechnung der Varianzkomponenten zeigen [Sieg und Hirsch \(2000a\)](#) auf.

#### 4.6.4 Alternative Schätzung der Varianzkomponenten

Koch (2004) konstatiert, dass bei ungünstig gewählten Näherungswerten für die Varianzkomponenten  $s_i^2$  die nach Gleichung (4.48) geschätzten Verbesserungsfaktoren negative Werte annehmen können. Alternativ ist das Berechnungsschema nach (Förstner 1979) zu verwenden:

$$\hat{\boldsymbol{\alpha}} = \mathbf{H}\mathbf{q} \quad (4.51)$$

mit

$$\mathbf{H} = [\text{diag}(1/\text{sp}(\mathbf{W}\mathbf{V}_i))] \quad \text{und} \quad \mathbf{q} = [\mathbf{1}^T \boldsymbol{\Sigma}_l^{-1} \mathbf{V}_i \boldsymbol{\Sigma}_l^{-1} \mathbf{1}]$$

Der Rechenaufwand zur Lösung dieser alternativen Schätzung ist im Vergleich zum Ansatz (4.48) geringer; allerdings konvergiert dieses Schema deutlich schlechter. Konsequenterweise ist die Verwendung von Lösungsweg (4.51) für die erste(n) Iteration(en) sinnvoll. Anschließend kann mit dem Lösungsweg (4.48) eine rasche Konvergenz erreicht und die Varianzen der Varianzkomponenten nach Gleichung (4.50) berechnet werden.

## 4.7 Random Sample Consensus

### 4.7.1 Motivation

Geodätische Ausgleichungsverfahren schätzen unbekannte Parameter einer Funktion bzw. eines geometrisch-physikalischen Modells für eine Reihe von Messwerten. Eine Ausgleichungsaufgabe ist eine mathematische Optimierungsmethode und fordert im Rahmen einer Ausgleichung nach z. B. vermittelnden Beobachtungen (Abschnitt 4.4.2), dass die Modellparameter so zu bestimmen sind, dass die „gewogene“ Summe der Fehlerquadrate minimiert wird. Die Schätzung eines Modells nach den Rechenvorschriften der MKQ basiert auf normalverteilten zufälligen Messabweichungen, welche auf alle Verbesserungen verteilt werden. Es wird dabei davon ausgegangen, dass das auszugleichende Datenmaterial frei von groben Fehlern ist. Befinden sich dennoch wenige Ausreißer im Datenmaterial, kommen Methoden der Ausreißerdetektion zum Einsatz (Abschnitt 4.5.4). Liegen jedoch viele grobe Fehler vor, versagen diese Ansätze. Bei einer ungünstigen Konstellation von Modell- und Datenfehlern können sogar fehlerfreie Beobachtungen verdächtigt und fehlerbehaftete Beobachtungen als fehlerfrei angesehen werden. Als Konsequenz konvergiert das Ausgleichungssystem bei einer Wiederholung bei einem falschen Parametersatz<sup>2</sup>.

### 4.7.2 Einsatzbereich

Der Einsatz des *Random Sample Consensus* (RANSAC; Fischler & Bolles 1981) als robustes Schätzverfahren bietet sich im Vorfeld einer Ausgleichungsrechnung an, um eine hohe Anzahl an Ausrei-

<sup>2</sup> Für diesen Abschnitt wird vereinbart, dass fehlerfreie Beobachtungen frei von groben und systematischen Fehlern sind; nicht aber frei von zufälligen Fehlern.

ßern im Beobachtungsmaterial zu detektieren und auszuschließen (Abbildung 4.4a). Solche groben Fehler werden gewöhnlich bereits während der Datenaufnahme durch Mehrfachmessungen bzw. bei der Datenaufbereitung durch Plausibilitätskontrollen ausgeschlossen (Abschnitt 4.5.1). Durch die in den letzten Jahrzehnten zugenommene Automatisierung von Messprozessen, vor allem im Bereich des maschinellen Sehens, erhöht sich der Umfang der gemessenen Daten erheblich. Eine interaktive Kontrolle der erhobenen Messwerte durch einen Operator ist in diesem Fall nicht praktikabel. Folglich erhöht sich auch der Ausreißeranteil.

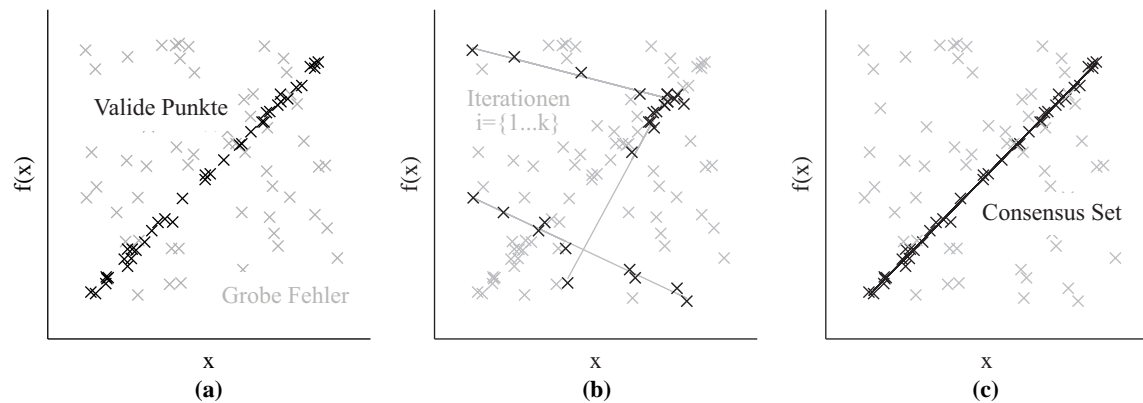
Mögliche Einsatzbereiche des RANSAC sind somit bei der Auswertung automatischer Messungen zu finden. Eine typisch photogrammetrische Anwendung ist z. B. die relative Orientierung eines Bildpaares anhand automatisch gemessener homologer Punkte in den Bildern. Die Lösung des Korrespondenzproblems ist aufgrund multipler Lösungen nicht trivial und kann im Messprozess zahlreiche Fehlkorrespondenzen verursachen (Abschnitt 7.3). RANSAC kann hier vor die ausgleichende Lösung der Orientierungsaufgabe geschaltet werden, um die aufgetretenen Fehlzuordnungen auszuschließen (Schulze 2010). Ein zweites Beispiel ist die Einpassung von geometrischen Primitiven wie Gerade, Ebene, Kugel oder Zylinder in eine 3D-Punktwolke, erzeugt mit z. B. einem terrestrischen Laserscanner (TLS; Beder & Förstner 2006) oder einer 3D-Kamera (Westfeld u. a. 2009). Der robuste Schätzer RANSAC kann hier trotz verrauschter Messwerte die Punktwolken zuverlässig in einzelne geometrische Körper modellgestützt segmentieren. Zur Bereitstellung grob fehlerfreier Beobachtungen bedient sich die in Kapitel 11 vorgestellte Kalibrieroutine ebenfalls *Random Sample Consensus* (RANSAC) als robustes Schätzverfahren.

### 4.7.3 Prinzip

Erstmals 1981 von Fischler und Bolles vorgestellt, verwendet der RANSAC-Algorithmus als iteratives Verfahren eine geschlossene Lösung und erweitert diese sukzessiv, falls weitere konsistente Daten vorliegen. Per Zufallsgenerator wird eine minimale Beobachtungsanzahl gewählt, die gesuchten Modellparameter daraus berechnet und die Abweichungen der verbleibenden Beobachtungen zum vorläufigen Modell analysiert (Abbildung 4.4b). Dieser Vorgang wird so lange wiederholt, bis definierte Abbruchkriterien wie eine maximale Iterationsanzahl oder eine Mindestmengenanzahl erreicht sind (Abbildung 4.4c).

Im Einzelnen sind die folgenden Schritte durchzuführen:

1. Bestimme zufällig eine Teilmenge  $S_i$  bestehend aus  $n$  Beobachtungen aus einem Datensatz der Größe  $N$  und berechne daraus eine direkte Lösung  $M_i$  der Modellparameter.
2. Erweitere die Teilmenge  $S_i$  um Beobachtungen, deren Abweichungen zum berechneten Modell  $M_i$  kleiner als eine gegebene Fehlerschranke  $\varepsilon$  sind.
3. Wiederhole die Schritte 1 und 2 bis die Größe der aktuellen Teilmenge  $S_i$  eine gegebene Mindestanzahl  $N_S$  überschreitet oder eine maximale Iterationsanzahl  $k$  erreicht ist.
4. Verwende die Beobachtungen der größten Teilmenge  $S_i$  zur statistischen Schätzung der Modellparameter. Nutze die zugehörigen Modellparameter  $M_i$  als Initialwerte.



**Abb. 4.4:** Prinzip des **RANSAC**-Algorithmus am Beispiel einer Geradeneinpassung: Ausgehend von einem Datensatz mit einem Ausreißeranteil von deutlich über 50 % (a) ermittelt **RANSAC** mit einer Wahrscheinlichkeit von  $w$  iterativ (b) nach maximal  $k$  Versuchen den optimalen Parametersatz (c).

Im Gegensatz zu statistischen Schätzverfahren wie der **MKQ** werden nicht alle Beobachtungen gemeinsam verwendet, um eine valide Teilmenge zu ermitteln. Vielmehr wird eine zunächst als fehlerfrei angenommene minimal notwendige Anzahl an Beobachtungen zufällig aus dem Datensatz gewählt. Erst durch die Unterstützung weiterer Beobachtungen wird diese Annahme bestätigt und das Modell akzeptiert. Diese Vorgehensweise ermöglicht für **RANSAC** einen Bruchpunkt von deutlich über 50 % und macht das Verfahren in Kombination mit einer nachgeschalteten Kleinsten-Quadrate-Anpassung zu einem vielseitigen und leistungsfähigen Werkzeug der Parameterschätzung.

#### 4.7.4 Steuerparameter

Das **RANSAC**-Verfahren wird im wesentlichen durch drei Parameter gesteuert:

##### Fehlertoleranz

Die Fehlertoleranz  $\varepsilon$  gibt den maximal zulässigen Abstand einer Beobachtung vom Modell an. Die Wahl dieses Parameters beeinflusst maßgeblich den Erfolg von **RANSAC**. Wird die Fehlerschranke  $\varepsilon$  zu groß gewählt, werden neben fehlerfreien Beobachtungen auch zahlreiche Ausreißer in die Teilmenge aufgenommen. Dies kann zur Akzeptanz eines eigentlich falschen Parametersatzes führen. Ist hingegen die Fehlerschranke  $\varepsilon$  zu klein angesetzt, unterstützen meist nicht alle eigentlich fehlerfreien Beobachtungen das Modell. Die Festlegung der Fehlertoleranz  $\varepsilon$  erfolgt empirisch oder über die bekannte a-priori Standardabweichung der Beobachtungen.

##### Iterationsanzahl

Die maximale Anzahl  $k$  an Iterationen gibt an, wie viele Teilmengen höchstens gezogen werden sollen. Die Iterationsanzahl lässt sich so festlegen, dass bei einer gegebenen Wahrscheinlichkeit  $z$

mindestens eine zufällig gezogene Stichprobe fehlerfrei ist:

$$k = \frac{\log(1 - z)}{\log(1 - w^n)} \quad (4.52)$$

Die Wahrscheinlichkeit  $w$ , dass die gewählte Beobachtung innerhalb einer gegebenen Fehlertoleranz  $\varepsilon$  liegt, kann auch durch das Verhältnis von fehlerfreien zu fehlerbehafteten Beobachtungen ausgedrückt werden. Der Ausreißeranteil ist in der Regel a-priori nicht bekannt. Er wird meist durch Erfahrungswerte eher zu hoch gesetzt und im Laufe der Berechnungen adaptiv nach unten korrigiert.

Ein typischer Wert für die Erfolgswahrscheinlichkeit  $z$  ist 99%. Für das in [Abbildung 4.4](#) gezeigte Beispiel einer Geradeneinpassung ergibt sich somit für  $n = 2$  Geradenparameter und einen Ausreißeranteil von ca. 60% bei insgesamt 175 Beobachtungen die maximale Iterationsanzahl zu  $k = 27$ . Bei gleicher Erfolgswahrscheinlichkeit erhöht sich  $k$  exponentiell in Abhängigkeit von der Minimal-konfiguration  $n$  und verschärft sich bei steigender Ausreißeranzahl. Die Iterationsanzahl ist somit nicht von der Gesamtanzahl  $N$  abhängig.

#### Größe der Teilmenge

Ein Schwellenwert  $N_S$  für die Größe einer zu akzeptierenden Teilmenge kann eingeführt werden, um den Algorithmus nach Erreichen vorzeitig abubrechen. Die Mindestgröße ist, wie auch der Ausreißeranteil, im Vorfeld nicht bekannt. Wurde der Ausreißeranteil bereits abgeschätzt, kann der Schwellenwert näherungsweise über  $N_S = w \cdot N$  berechnet werden.

#### 4.7.5 Weiterentwicklungen

Der **RANSAC**-Algorithmus wurde in den letzten Jahren durch verschiedene Ansätze weiterentwickelt. Der Vollständigkeit halber werden an dieser Stelle drei wichtige Erweiterungen kurz vorgestellt. [Torr und Zisserman \(2000\)](#) kompensieren das Problem einer a-priori meist nicht genau bekannten Fehlerschranke  $\varepsilon$  durch die Einführung von gewichteten Fehlern für die Beobachtungen. Diese Weiterentwicklung wird als *M-Estimator Sample Consensus* (MSAC) bezeichnet. [Chum u. a. \(2003\)](#) erweitern das Verfahren zu einem *Local Optimised Random Sample Consensus* (LO-RANSAC) und bestimmen nach jedem zweiten Schritt ein optimiertes Modell und somit eine verbesserte Teilmenge  $S_i^*$  mittels eines Ausgleichsverfahrens. *GASAC* (*Genetic Algorithm Sampling Consensus*; [Rodehorst & Hellwich 2006](#)) modifiziert den ursprünglichen Algorithmus in der Art, dass mehrere Minimallösungen berechnet werden und die zufällige Stichprobenziehung durch ein weitgehend systematisches Probieren verbessert wird. Dies führt zu einer effizienteren und schnelleren Berechnung des optimalen Parametersatzes.





## Photogrammetrische und geodätische Grundlagen

Dieses Kapitel behandelt kurz photogrammetrische Definitionen, Methoden und Verfahren. Sie werden für das Verständnis der nachfolgenden Ausführungen vorausgesetzt.

Die in der Nahbereichsphotogrammetrie üblichen Aufnahmekameras bilden räumliche Objekte zentralperspektiv auf der Bildebene ab. Das mathematische Modell der Zentralprojektion des Raums sowie die sich daraus ergebenden Zusammenhänge zwischen Bild-, Kamera- und Objektkoordinatensystemen werden im [Abschnitt 5.1](#) behandelt. Die Photogrammetrie ist ein dreidimensionales Messverfahren zur Rekonstruktion der Form und der Lage beliebiger Objekte aus z. B. Fotografien. Die Methoden einer solchen Objektrekonstruktion diskutiert der [Abschnitt 5.2](#).

Die Bestimmung der exakten Kamerageometrie sowie die Orientierung eines Einzelbilds oder eines Mehrbildverbands im Raum sind Gegenstand der [Abschnitte 5.3](#) und [5.4](#).

### 5.1 Abbildungsmodell

#### 5.1.1 Zentralprojektion des Raumes

Die mathematische Grundlage eines konventionellen Abbildungsvorganges in der Photogrammetrie ist das Modell der Zentralprojektion. Bei einer Zentralprojektion wird, entsprechend dem als bekannt vorausgesetztem geometrischen Prinzip einer Lochkamera, ein dreidimensionaler Objektpunkt  $\mathbf{P}(X, Y, Z)$  auf der Sensorebene als Bildpunkt  $\mathbf{P}'(x', y')$  abgebildet ([Abbildung 5.1](#)). Der Projektionsstrahl durchläuft dabei das Projektionszentrum  $\mathbf{X}_0(X_0, Y_0, Z_0)$  der Kamera; Bildpunkt, Projektionszentrum und korrespondierender Objektpunkt sind somit kollinear ([Abschnitt 5.1.5](#)). Unter Kenntnis der Abbildungsgeometrie des Sensors (innere Orientierung und Zusatzparameter; [Abschnitt 5.1.3](#)) und der Lage der Kamera im Raum (äußere Orientierung; [Abschnitt 5.1.4](#)) können die Form und die Lage eines Objekts über die Rekonstruktion von Strahlenbündeln bestimmt werden.

#### 5.1.2 Koordinatensysteme

Die Messung der ursprünglichen Beobachtungen wird in einem zweidimensionalen kartesischen Bildkoordinatensystem  $x'y'$  durchgeführt, welches der Festlegung eines kamerafesten Bezugs dient. Mit der Erweiterung um eine senkrecht auf der Bildebene stehenden  $z'$ -Achse gelingt der Übergang vom ebenen Messbild ins Dreidimensionale. Der Ursprung dieses neu definierten  $x'y'z'$ -Kamerakoordinatensystems wird von der Bildmitte in das Projektionszentrum verschoben. Schließlich werden in einem dreidimensionalen kartesischen  $XYZ$ -Objektkoordinatensystem die Punkte eines Objekts sowie dessen Orientierung übergeordnet festgelegt.

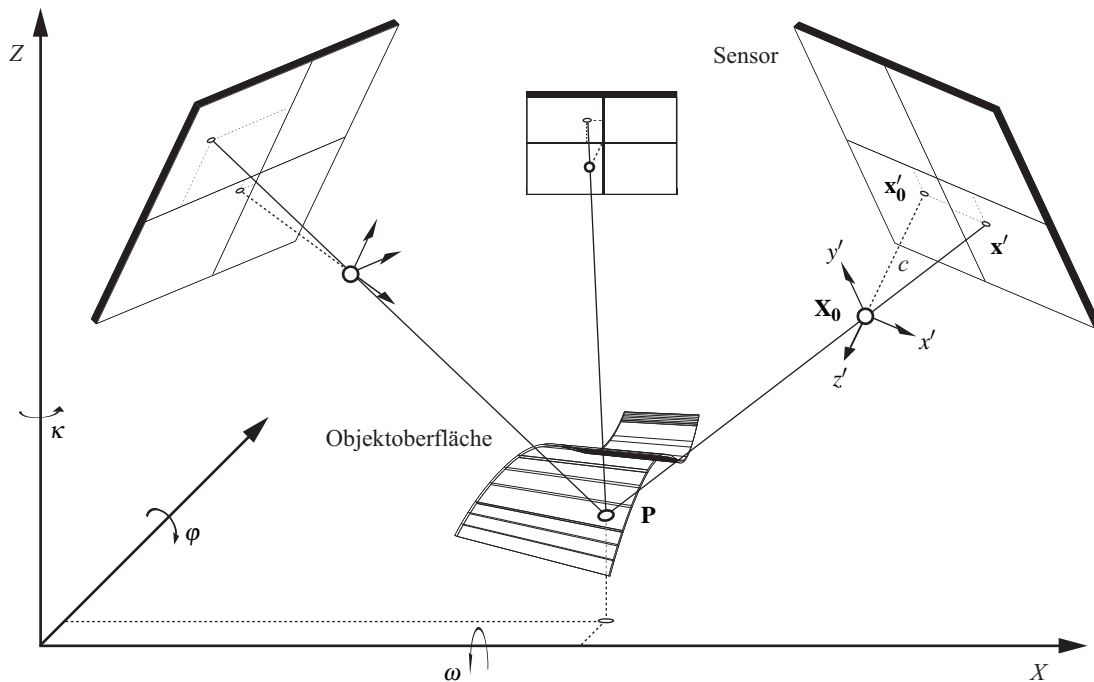


Abb. 5.1: Photogrammetrisches Messprinzip.

Die **Abbildung 5.1** stellt den Zusammenhang zwischen dem Bild-, Kamera- und Objektkoordinatensystem grafisch dar.

### 5.1.3 Innere Orientierung und Zusatzparameter

Die innere Orientierung einer Kamera beschreibt den Zusammenhang zwischen dem Bildkoordinatensystem  $x'y'$  und dem Kamerakoordinatensystem  $x'_0y'_0z'$  (**Abbildung 5.1**). Sie wird durch die Lage des Projektionszentrums  $\mathbf{X}_0$  in Bezug zur Bildebene sowie etwaigen Abweichungen vom Idealmodell einer Zentralprojektion definiert. Die Parameter der inneren Orientierung sind kameraspezifische Größen. Zu ihnen gehören die Kamerakonstante  $c$ , welche den lotrechten Abstand zwischen der Bildebene und dem Projektionszentrum quantifiziert, sowie der Bildhauptpunkt  $\mathbf{x}'_0(x'_0, y'_0)$  als Durchstoßpunkt der optischen Achse durch die Bildebene. Für einen im Bild gemessenen Punkt  $\mathbf{P}(x', y')$  lautet der zugehörige Abbildungsvektor:

$$\mathbf{x}' = \begin{bmatrix} x' - x'_0 - \Delta x' & y' - y'_0 - \Delta y' & -c \end{bmatrix}^T \quad (5.1)$$

Darüber hinaus können Zusatzparameter von Bildfehler beschreibenden Funktionen eingeführt werden. Die achsbezogenen Korrekturwerte  $\Delta \mathbf{x}'(\Delta x', \Delta y')$  in Gleichung (5.1) modellieren mögliche Abweichungen vom zentralperspektivischen Modell, verursacht durch u. a. Fehler der Optik (z. B. Objektivverzeichnung), topografische Fehler des Sensors oder Fehler der Kameramechanik (z. B. Drift der Sensorfläche im Kameragehäuse). Die Objektivverzeichnung hat den größten Einfluss (**Luh-**

mann 2003) und wird in der Nahbereichsphotogrammetrie häufig durch den Zusatzparametersatz von Brown (1971) ausgedrückt.

Drei Koeffizienten eines Polynoms modellieren den radial-symmetrischen Einfluss:

$$\Delta r'_{rad} = A_1 \cdot r' (r'^2 - r_0'^2) + A_2 \cdot r' (r'^4 - r_0'^4) + A_3 \cdot r' (r'^6 - r_0'^6) \quad (5.2)$$

Der Bildradius  $r'$  bezeichnet den Abstand eines gemessenen Bildpunkts vom Bildhauptpunkt. Der konstante und zumeist mit  $2/3$  des maximalen Bildradius initialisierte Parameter  $r_0$  bewirkt einen zweiten Nulldurchgang der Verzeichnungskurve, wodurch die minimalen und maximalen Verzeichnungswerte gleich groß werden (*balanced radial distortion*). Die Korrektur der Bildkoordinaten erfolgt sodann proportional:

$$\Delta \mathbf{x}'_{rad} = \mathbf{x}' \cdot \frac{\Delta r'_{rad}}{r'} \quad (5.3)$$

Die radial-asyymetrische und tangentielle Verzeichnung (Dezentrierverzeichnung) wird durch zwei weitere Koeffizienten beschrieben:

$$\begin{aligned} \Delta x'_{tan} &= B_1 \cdot (r'^2 + 2x'^2) + 2B_2 \cdot x' \cdot y' \\ \Delta y'_{tan} &= B_2 \cdot (r'^2 + 2y'^2) + 2B_1 \cdot x' \cdot y' \end{aligned} \quad (5.4)$$

In der digitalen Nahbereichsphotogrammetrie wird der Brownsche Parametersatz (1971) um zwei Affintransformationsparameter erweitert, welche die Abweichungen des Bildkoordinatensystems von Orthogonalität und Gleichmaßstäbigkeit der Koordinatenachsen modellieren. Diese Affinitäts- und Scherungseinflüsse werden durch Abweichungen der Sensorelemente von ihrer Sollgröße und durch Fehler bei der Digitalisierung des analogen Bildsignals verursacht. Nach z. B. El-Hakim (1986) werden sie wie folgt korrigiert:

$$\begin{aligned} \Delta x'_{aff} &= C_1 \cdot x' + C_2 \cdot y' \\ \Delta y'_{aff} &= 0 \end{aligned} \quad (5.5)$$

Zusammengefasst ergibt sich die Gesamtkorrektur der auftretenden Abbildungsfehler zu:

$$\Delta \mathbf{x}' = \Delta \mathbf{x}'_{rad} + \Delta \mathbf{x}'_{tan} + \Delta \mathbf{x}'_{aff} \quad (5.6)$$

Die Bestimmung der inneren Orientierungselemente geschieht im Rahmen einer Kamerakalibrierung im Allgemeinen indirekt aus photogrammetrischen Beobachtungen. Sie ist bei jeder Änderung der aktuellen Blendeneinstellung und Fokussierung sowie bei einem Objektivwechsel zu wiederholen. Nach Maas (1997) werden die Verfahren zur Kalibrierung einer Messkamera in der Photogrammetrie häufig klassifiziert in Laborkalibrierung, Testfeldkalibrierung (Abschnitt 5.3.3) sowie Selbst- und Simultankalibrierung, teils mit geodätischen Zusatzinformationen (Abschnitt 5.4.3).

### 5.1.4 Äußere Orientierung

Als äußere Orientierung wird die räumliche Transformation des  $x'y'z'$ -Kamerakoordinatensystems in das übergeordnete  $XYZ$ -Objektkoordinatensystem bezeichnet. Entsprechend der Transformationsvorschrift (4.4) legen die Parameter der äußeren Orientierung die räumliche Lage und die Ausrichtung der Kamera im Zielsystem durch drei Translationen und drei Rotationen fest (Abbildung 5.1). Das Projektionszentrum  $\mathbf{X}_0$  der Kamera als gemeinsamer Bezugspunkt der Systeme bzw. der zugehörige Translationsvektor legen die räumliche Lagerung des Bildkoordinatensystems im Objektkoordinatensystem fest. Die Verdrehung der Systeme zueinander wird in der terrestrischen Photogrammetrie bei nahezu horizontaler Aufnahme durch die drei Drehwinkel  $\omega$  (Azimut),  $\varphi$  (Neigung um Kippachse) und  $\kappa$  (Kantung um Aufnahmeachse) beschrieben.

Die Parameter der äußeren Orientierung werden indirekt durch die Messung der Bildkoordinaten von mindestens drei bekannten Objektpunkten ermittelt (Orientierung eines Einzelbilds; Abschnitt 5.3). Anschließend ist die Transformation eines Bildkoordinatenvektors  $\mathbf{x}'$  in einen absolut im Raum definierten Strahl möglich.

### 5.1.5 Kollinearitätsgleichungen

Ausgehend von der im Abschnitt 4.2 beschriebenen räumlichen Koordinatentransformation lautet die Abbildung eines Bildpunkts in den Objektraum:

$$\mathbf{X} = \mathbf{X}_0 + m \cdot \mathbf{R} \cdot \mathbf{x}'$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + m \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} x' - x'_0 - \Delta x' \\ y' - y'_0 - \Delta y' \\ -c \end{bmatrix} \quad (5.7)$$

Der Maßstabsfaktor  $m$  ist unbekannt und für jeden Objektpunkt verschieden. Aus einem Einzelbild kann folglich nur der Richtungsvektor vom einem Bildpunkt über das Projektionszentrum zum gesuchten Objektpunkt bestimmt werden, nicht aber dessen räumliche Lage. Die Rücktransformation in den Bildraum ist hingegen eindeutig:

$$\mathbf{x}' - \mathbf{x}'_0 - \Delta \mathbf{x}' = \frac{1}{m} \cdot \mathbf{R}^{-1} \cdot \mathbf{X} - \mathbf{X}_0$$

$$\begin{bmatrix} x' - x'_0 - \Delta x' \\ y' - y'_0 - \Delta y' \\ -c \end{bmatrix} = \frac{1}{m} \cdot \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \quad (5.8)$$

Werden die beiden ersten Zeilen des Gleichungssystems (5.8) durch die dritte dividiert, eliminiert sich der unbekannte Maßstabsfaktor  $m$ . Es folgen die Kollinearitätsgleichungen, welche das Modell

der Zentralprojektion des Raums mathematisch beschreiben (Abschnitt 5.1.1):

$$\begin{aligned} x' &= x'_0 - c \cdot \frac{r_{11} \cdot (X - X_0) + r_{21} \cdot (Y - Y_0) + r_{31} \cdot (Z - Z_0)}{r_{13} \cdot (X - X_0) + r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0)} + \Delta x' \\ y' &= y'_0 - c \cdot \frac{r_{12} \cdot (X - X_0) + r_{22} \cdot (Y - Y_0) + r_{32} \cdot (Z - Z_0)}{r_{13} \cdot (X - X_0) + r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0)} + \Delta y' \end{aligned} \quad (5.9)$$

Die Kollinearitätsgleichungen gelten als die Grundgleichungen der Photogrammetrie und finden bei der Objektrekonstruktion bspw. beim räumlichen Vorwärts- und Rückwärtseinschneiden (Abschnitte 5.2 und 5.3.1) oder bei der Bündeltriangulation (Abschnitt 5.2) Anwendung.

## 5.2 Objektrekonstruktion

Bei einem räumlichen Vorwärtsschnitt werden die 3D-Koordinaten  $\mathbf{X}$  eines Objektpunkts aus seinen Bildkoordinaten  $\mathbf{x}'$  und den inneren und äußeren Orientierungsparametern entsprechend dem zentralperspektivischen Abbildungsmodell (5.7) berechnet. Dabei kommen Verfahren der Einzel-, Stereo- oder Mehrbildauswertung zum Einsatz.

Bereits im Abschnitt 5.1.5 wurde konstatiert, dass aus einer reinen Einzelbildkoordinatenmessung lediglich die Richtung zum Objektpunkt, nicht aber dessen absolute Lage bestimmt werden kann. Die Auflösung der Kollinearitätsgleichungen (5.9) nach den Objektkoordinaten ergäbe die Objektlage  $(X, Y)$  als Funktion der Tiefenkoordinate  $Z$ ; einem Bildpunkt können somit entlang des Projektionsstrahls unendlich viele Objektpunkte zugeordnet werden. Erst die Integration geometrischer Zusatzinformationen, bspw. durch den Schnitt des Projektionsstrahls mit einer Ebene (Projektivtransformation) oder die Verwendung dreidimensionaler Objektmodelle (Interpolation der Tiefenkomponente aus einem digitalen Oberflächenmodell), erlaubt die Objektrekonstruktion aus einem Einzelbild (z. B. Luhmann 2003).

Die räumliche Lage eines Objekts kann genau dann eindeutig und ohne Objektwissen bestimmt werden, wenn aus mehreren Bildern rekonstruierte Strahlenbündel im Objektraum zum Schnitt gebracht werden. Bei hergestellter innerer und äußerer Orientierung liefert bereits die Messung homologer Bildpunkte in zwei von unterschiedlichen Standpunkten und aus möglichst konvergenten Richtungen akquirierten Kameraaufnahmen vier Beobachtungsgleichungen (5.9) zur Berechnung der drei unbekanntenen Objektpunktkoordinaten. Die Konfiguration kann durch das Hinzufügen beliebig vieler orientierter Objektaufnahmen zu einem Mehrbildverband erweitert werden. Die unbekanntenen Objektkoordinaten werden im Zuge einer Ausgleichung nach vermittelnden Beobachtungen berechnet (Abschnitt 4.4.2).

## 5.3 Einzelbildorientierung

Die Orientierung eines Einzelbilds beschränkt sich zunächst auf die Herstellung der äußeren Orientierung, also der indirekten Berechnung der Koordinaten des Projektionszentrums  $\mathbf{X}_0$  sowie der

Drehwinkel ( $\omega, \varphi, \kappa$ ) über photogrammetrische Bildpunktmessungen bekannter Passpunkte. Die dabei üblichen Rechenverfahren basieren entweder auf den Kollinearitätsgleichungen (Abschnitt 5.3.1) oder auf projektiven Beziehungen (Abschnitt 5.3.2) und sollen im Folgenden kurz vorgestellt werden.

### 5.3.1 Räumlicher Rückwärtsschnitt

Bei einem räumlichen Rückwärtsschnitt wird das durch die bekannten Objektpunkte und das Projektionszentrum aufgespannte Strahlenbündel rechnerisch eindeutig in die korrespondierenden Bildpunkte eingepasst und daraus die räumliche Lage und Orientierung eines einzelnen Bilds rekonstruiert. Bei einer bekannten Kamerageometrie müssen die Bildkoordinaten von mindestens drei nicht auf einer Geraden liegenden Passpunkten gemessen werden. Wurde die innere Orientierung noch nicht hergestellt, erhöht sich die Anzahl auf fünf räumlich verteilte Punkte. Der räumliche Rückwärtsschnitt wird häufig als eine Ausgleichung nach vermittelnden Beobachtungen formuliert (Abschnitt 4.4.2). Die Bildkoordinaten gehen als Beobachtungen in die im Abschnitt 5.3.1 hergeleiteten Kollinearitätsgleichungen (5.9) ein. Diese sind durch die partiellen Ableitungen der Unbekannten an den Stellen der Näherungswerte zu linearisieren und im weiteren Verlauf zum Aufbau einer Koeffizientenmatrix sowie eines verkürzten Beobachtungsvektors heranzuziehen. Die Näherungswerte der unbekannt orientierungselemente werden im Gauß-Markov-Modell (GMM) schließlich iterativ verbessert.

Prinzipbedingt gehen bei einer Einzelbildorientierung die Objektpunktkoordinaten als Konstanten in die Ausgleichung ein. Die Genauigkeit dieser zuvor gemessenen Passpunkte kann somit die Systemgenauigkeit limitieren. Die notwendigen Näherungswerte stammen z. B. aus im Vorfeld getätigten geodätischen Messungen. Eine elegante und flexible Alternative ist die Vorschaltung eines linearen Berechnungsverfahrens, welches ohne Näherungswerte arbeitet (Abschnitt 5.3.2). Der räumliche Rückwärtsschnitt als ein Verfahren zur Orientierung eines Einzelbilds kann um prinzipiell beliebig viele Aufnahmen zu einem Mehrbildverband erweitert werden (Bündelblockausgleichung mit festen Objektpunktkoordinaten; z. B. Schneider 2009).

### 5.3.2 Direkte lineare Transformation

Das bekannteste lineare Berechnungsverfahren ist die direkte lineare Transformation (DLT; Abdel-Aziz & Karara 1971). Sie bestimmt die Orientierungsdaten eines Einzelbilds mit insgesamt elf Transformationsparametern:

$$x' = \frac{L_1X + L_2Y + L_3Z + L_4}{L_9X + L_{10}Y + L_{11}Z + 1} \quad y' = \frac{L_5X + L_6Y + L_7Z + L_8}{L_9X + L_{10}Y + L_{11}Z + 1} \quad (5.10)$$

Durch eine einfache Umstellung der Transformationsgleichungen (5.10) wird ein linearer Zusammenhang zwischen den gemessenen Bildkoordinaten und den zu bestimmenden Objektkoordinaten

hergestellt:

$$\begin{aligned} L_1X + L_2Y + L_3Z + L_4 - x'L_9X - x'L_{10}Y - x'L_{11}Z - x' &= 0 \\ L_5X + L_6Y + L_7Z + L_8 - y'L_9X - y'L_{10}Y - y'L_{11}Z - y' &= 0 \end{aligned} \quad (5.11)$$

Das Gleichungssystem wird als eine reguläre Ausgleichungsaufgabe im **GMM** formuliert und ohne die Notwendigkeit von Näherungswerten gelöst (**Abschnitt 4.4.2**).

Dem Vorteil eines linearen Gleichungssystems stehen jedoch auch Nachteile gegenüber: Die **DLT** benötigt mindestens sechs Passpunkte, welche nicht in einer Ebene liegen dürfen. Auch führen bestimmte Aufnahmekonstellationen zu schlecht konditionierten oder gar singulären Gleichungssystemen. Bekannte innere Orientierungsparameter können nicht in das Transformationsmodell eingeführt werden, da andernfalls das Gleichungssystem überparametrisiert ist.

Der Ansatz der **DLT** wird oft zur Beschaffung von Näherungswerten für die Parameter der äußeren Orientierung herangezogen. So auch im Rahmen einer integrierten Kalibrierroutine, welche im **Kapitel 11** vorgestellt wird. Für eine Darstellung der Beziehungen zwischen den Parametern der inneren als auch äußeren Orientierung und den **DLT**-Parametern sei an dieser Stelle auf z. B. (**Luhmann 2003**) verwiesen.

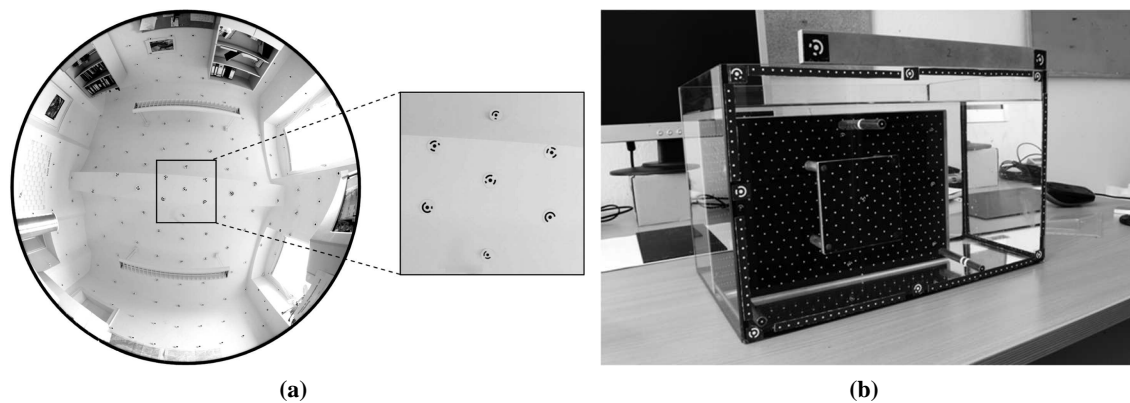
### 5.3.3 Referenzfeldkalibrierung

Für die Orientierung eines Einzelbilds wurde in den **Abschnitten 5.1.5, 5.3.1** und **5.3.2** das zentralperspektivische Abbildungsmodell mathematisch aufgestellt und die darauf aufbauenden Berechnungsverfahren erläutert. Die folgenden Ausführungen befassen sich kurz mit der praktischen Umsetzung, also der Kamerakalibrierung.

Die Aufnahme eines signalisierten Referenzfelds stellt eine einfache und robuste Methode der Kamerakalibrierung dar (**Abbildung 5.2**). Liegen die Koordinaten von ausreichend vielen und räumlich verteilten Referenzpunkten mit übergeordneter Genauigkeit vor, genügt ein räumlicher Rückwärtschnitt im Bündel zur Herstellung der Kameraorientierung. Einschränkungen ergeben sich in der Praxis bei der Anfertigung eines stabilen, möglichst transportablen und in seiner Ausdehnung der späteren Anwendung angepassten Referenzfelds ebenso wie bei der Bereitstellung hinreichend genauer Passpunktkoordinaten (**Maas 1997**).

Die Kalibrieranordnung muss bei einem ebenen Referenzfeld um mehrere schräg gerichtete Aufnahmen, bei nur näherungsweise bekannten Objektpunktkoordinaten zusätzlich um verkantete Frontaufnahmen erweitert werden. Diese Mehrbildtriangulation soll Gegenstand des **Abschnitts 5.4** sein.





**Abb. 5.2:** Referenzfeldkalibrierung: (a) Hemisphärische Fischaugenobjektivaufnahme eines fest im Raum installierten Referenzfelds (Schneider 2009). (b) Aufnahme einer in einem Wasserbecken platzierten transportablen Referenzplatte zur Kamerakalibrierung in einer Mehrmedienumgebung (Mulsow 2010).

## 5.4 Mehrbildtriangulation

### 5.4.1 Übersicht

Die Bündelblockausgleichung (auch Mehrbildtriangulation) ist ein leistungsfähiges und genaues Verfahren zur gleichzeitigen Bildorientierung und Objektpunktbestimmung, welches in der (analytischen) Photogrammetrie bereits seit Ende der 70er Jahre existiert (Brown 1976). Auf der Grundlage des zentralperspektivischen Abbildungsmodells (5.9) ermöglicht die Kombination aus einem räumlichen Vorwärts- und Rückwärtsschnitt zum einen die Orientierung beliebig vieler im Raum angeordneter Bilder, zum anderen die vollständige dreidimensionale Rekonstruktion des aufgenommenen Objekts. Als Beobachtungen dienen einzig photogrammetrische Informationen in Form von Bildkoordinatenmessungen. Die Parameter der inneren und äußeren Orientierung sowie die zur Bildpunktmessung korrespondierenden Objektpunkte werden als Unbekannte eingeführt.

Das aufgestellte Gleichungssystem einer Bündelblockausgleichung ist in der Regel hoch redundant und stellt folglich ein Optimierungsproblem dar. Unter der Forderung eines optimalen Schnitts aller homologen Bildstrahlen in ihrem Objektpunkt wird es nach den Ausführungen des Abschnitts 4.4 ausgleichend gelöst. Neben den Orientierungselementen der Bilder, den 3D-Objektpunktkoordinaten und zusätzlich eingeführter Modellparametern können statistische Informationen über die Genauigkeit und Zuverlässigkeit des Gesamtmodells als auch der einzelnen geschätzten Modellparameter berechnet werden.

Die Anwendung von Mehrbildtechniken im Rahmen einer Bündelblockausgleichung ermöglicht die Überbrückung passpunktloser Räume, die Steigerung der Genauigkeit und Zuverlässigkeit sowie die Bestimmbarkeit von Parametern. Darüber hinaus ist die Bündelblockausgleichung flexibel in der Aufnahmeanordnung, der Wahl des Aufnahmesystems und der Aufnahmerichtung (Luhmann 2003).



In Anwendungen der digitalen Nahbereichsphotogrammetrie und darüber hinaus spielt sie bis heute eine große Rolle.

#### 5.4.2 Datumsfestlegung

Das aufgebaute photogrammetrische Netz ist ohne den Bezug zu einem übergeordneten Objektkoordinatensystem lediglich in seiner Form, nicht aber in seiner absoluten Lage und Orientierung im Raum festgelegt. Das im Zuge der Ausgleichsrechnung zu lösende Normalgleichungssystem besitzt somit einen Rang- bzw. Datumsdefekt, welcher nur durch eine entsprechende absolute Netzlagerung behoben werden kann.

Ein Objektkoordinatensystem wird eindeutig über die Festlegung des Ursprungs (3 Translationen), der Orientierung (3 Rotationen) und eines Maßstabs definiert. Diese insgesamt sieben Freiheitsgrade können über drei Passpunkte mit mindestens sieben Koordinateninformationen beschrieben und zur zwangsfreien Einpassung verwendet werden (minimale Datumsfestlegung). Diese sollten räumlich gut über den gesamten Mehrbildverband verteilt sein. In der Regel liegen in den Anwendungen der Nahbereichsphotogrammetrie mehr Passpunkte vor. Um Netzspannungen zu vermeiden, werden sie meist nicht als fehlerfreie Referenzpunkte sondern als beobachtete Größen mit einem ihrer Messgenauigkeit entsprechenden Gewicht in die Bündelblockausgleichung eingeführt und im Zuge der Optimierungsaufgabe verbessert.

Liegen keine Passpunktinformationen vor, kann das Datum auch über eine freie Netzausgleichung durch Auffelderung auf die Näherungskoodinaten  $\mathbf{X}^0$  aller unbekanntenen Objektpunkte (Neupunkte und Projektionszentren) festgelegt werden. Der Punkthaufen  $\mathbf{X}^0$  kann ohne geometrischen Zwang durch Verschiebung und Drehung in alle Koordinatenrichtungen und eine Systemskalierung in die im Rahmen einer z. B. vermittelnden Ausgleichung zu schätzenden Objektpunkte transformiert werden. Hierfür notwendig ist die Aufstellung von sieben linear unabhängigen Bedingungsgleichungen, nämlich drei für die Systemtranslation

$$\begin{array}{ccccccc} dX_1 & & +dX_2 & & + \dots & = & 0 \\ & dY_1 & & +dY_2 & & + \dots & = & 0 \\ & & dZ_1 & & +dZ_2 & + \dots & = & 0 \end{array} \quad (5.12)$$

drei für die differentielle Drehung des Koordinatensystems

$$\begin{array}{cccccccc} & -Z_1^0 dY_1 & +Y_1^0 dZ_1 & & -Z_2^0 dY_2 & +Y_2^0 dZ_2 & + \dots & = & 0 \\ Z_1^0 dX_1 & & -X_1^0 dZ_1 & +Z_2^0 dX_2 & & -X_2^0 dZ_2 & + \dots & = & 0 \\ -Y_1^0 dX_1 & +X_1^0 dY_1 & & -Y_2^0 dX_2 & +X_2^0 dY_2 & & + \dots & = & 0 \end{array} \quad (5.13)$$

sowie eine Maßstabsfestlegung

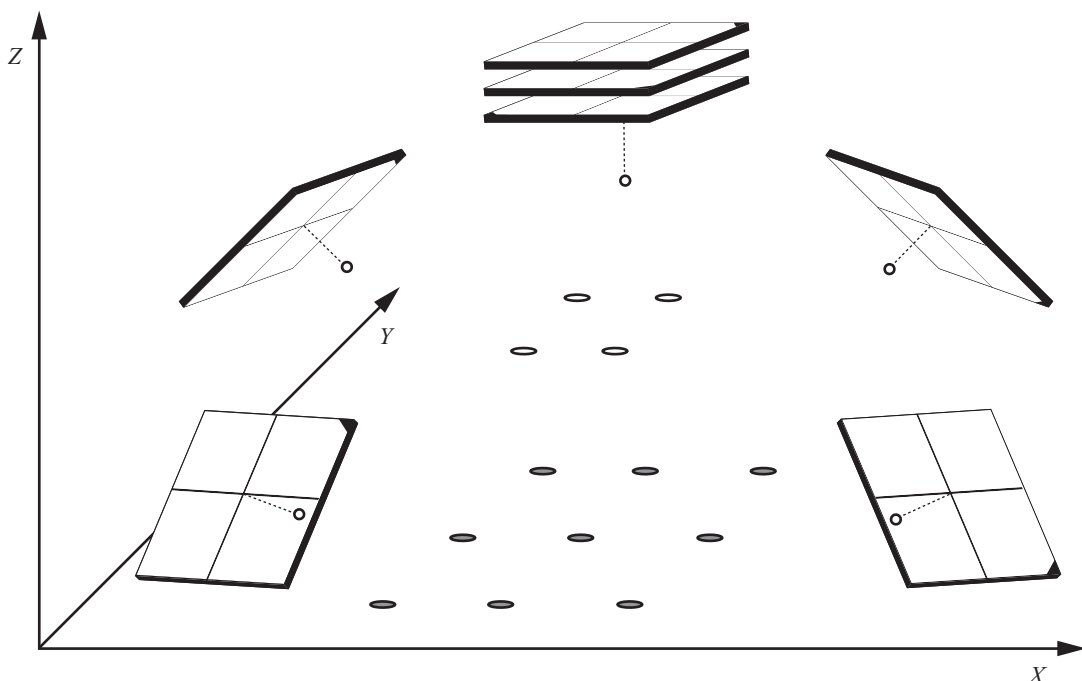
$$X_1^0 dX_1 + Y_1^0 dY_1 + Z_1^0 dZ_1 + X_2^0 dX_2 + Y_2^0 dY_2 + Z_2^0 dZ_2 + \dots = 0 \quad (5.14)$$

welche nach [Abschnitt 4.4.3](#) als zusätzliche Bedingungsgleichungen in das funktionale Modell integriert werden.

### 5.4.3 Selbst- und Simultankalibrierung

Eine Selbstkalibrierung bezeichnet die gleichzeitige Bestimmung der inneren und äußeren Orientierungsparameter der Kamera sowie möglicher Zusatzparameter ohne Objektrauminformation allein durch die Rekonstruktion von Strahlenbündeln aus photogrammetrischen Bildpunktmessungen. Eine hierfür geeignete Aufnahmekonfiguration zeigt [Abbildung 5.3](#). Insgesamt genügen sieben frontal und konvergent auf ein räumliches und in seinen Koordinaten nur näherungsweise bekanntes Referenzpunktfeld gerichtete Aufnahmen. Die vier konvergenten Aufnahmen ermöglichen eine genaue und zuverlässige Neupunktbestimmung unter einer guten Strahlenschnittgeometrie. Die um jeweils  $90^\circ$  gewälzten Frontalaufnahmen dienen der Bestimmbarkeit der Kameraparameter. Die Einhaltung der Rotationsstrategie vermeidet dabei Korrelationen zwischen den Parametern der inneren und äußeren Orientierung. Die Datumsfestlegung erfolgt meist durch eine freie Netzausgleichung [Abschnitt 5.4.2](#). Zur Stärkung der Netzgeometrie können geodätische Zusatzinformationen in Form von bspw. Strecken- und Richtungsmessungen oder auch durch parametrisierte geometrische Elemente (z. B. Gerade, Ebene, rotationssymmetrischer Körper) in die Bündelblockausgleichung integriert werden.

Bei einer Simultankalibrierung wird die Kamera in der Anwendung kalibriert und das Referenzfeld durch das Messobjekt ersetzt.



**Abb. 5.3:** Aufnahmekonfiguration zur Selbstkalibrierung (nach [Godding 1993](#)).

#### 5.4.4 Kombinierte Ausgleichung unterschiedlicher Beobachtungen

Die auf der **Methode der kleinsten Quadrate (MKQ)** basierende Bündelblockausgleichung ist bestens geeignet, um verschiedenartige Beobachtungstypen eines oder mehrerer Messgeräte zu berücksichtigen. Die Formulierung der unterschiedlichen geometrischen Modelle in einem funktionalen Zusammenhang sowie die Berücksichtigung der verschiedenen Messgenauigkeiten in Form von für das stochastische Modell aufzustellenden Gewichten erlaubt eine optimale Verknüpfung der Beobachtungen aller Messgeräte sowie der (gemeinsamen) unbekannt Parameter.

Bereits 1993 integrieren **Ackermann und Schade** GPS-Beobachtungen (**Globales Positionierungssystem**) zur unterstützenden Berechnung der absoluten Positionierung von Aufnahmen in einem Luftbildverband. Eine weitere Anwendung für die kombinierte Ausgleichung unterschiedlicher Beobachtungen liefert bspw. **Schneider (2009)**. Er entwickelt eine integrierte Methode zur Berechnung diskreter Objektpunkte, abgeleitet aus Beobachtungen eines terrestrischen Laserscanners und photogrammetrischer Bilddaten. Auch in dieser Arbeit werden unterschiedliche Beobachtungsarten in Kalibrierungs- und Bildanalyseanwendungen integriert (**Kapitel 11 und 14**).



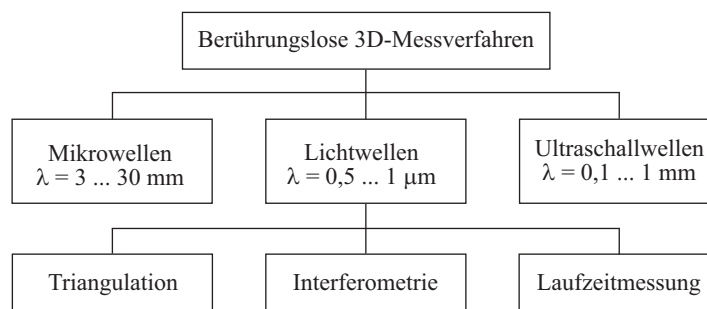
## Verfahren zur 3D-Objekterfassung

Das Kapitel behandelt verschiedene berührungslose Aufnahmeverfahren, um entweder die Objekt-oberfläche in Form eines Tiefenbilds oder das Objekt vollständig als Volumenbild zu rekonstruieren. Dabei werden die relevantesten Techniken zunächst kategorisiert und anschließend deren Prinzipien erläutert. Im Hinblick auf einen Vergleich mit einer neuartigen 3D-Kamera als monosensorielles System zur Generierung dynamischer Oberflächenmodelle (Kapitel 8) werden die Potentiale der jeweiligen Technik herausgearbeitet und Verfahrensgrenzen diskutiert.

Primär werden hier Oberflächenmessverfahren vorgestellt, entsprechend ihrer Relevanz teils prägnant oder ausführlicher. Sie können sowohl in direkter Konkurrenz zu 3D-Kameras stehen oder als zur Tiefenbildmesstechnik komplementäre Verfahren eingesetzt werden. Einzig das Silhouettenverfahren ist ein Vertreter der volumetrischen Rekonstruktionstechnik. Es wird in jüngster Zeit häufig mit distanzmessenden Kameras fusioniert und folglich an dieser Stelle aufgeführt.

### 6.1 Übersicht

Berührungslose Messverfahren zur 3D-Erfassung von Objekten werden nach Schwarte u. a. (1999) zunächst hinsichtlich ihres Informationsträgers in Lichtwellen-, Mikrowellen- und Ultraschallwellenverfahren unterschieden (Abbildung 6.1). In der photogrammetrischen Messtechnik kommen in der Regel Lichtwellenverfahren zum Einsatz.



**Abb. 6.1:** Klassifikation berührungsloser 3D-Messverfahren (nach Schwarte u. a. 1999).

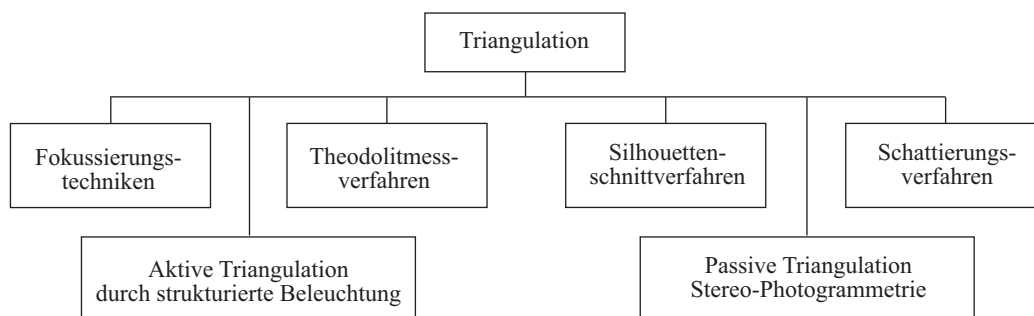
Nach Abbildung 6.1 ist eine weitere Kategorisierung für berührungslose 3D-Messverfahren auf der Basis von Lichtwellen möglich, wobei nun anwendungsspezifisch die Objektgröße und die geometrische Genauigkeit (Auflösung) mit berücksichtigt werden müssen. Die Triangulation ist ein Verfahren zur elektrooptischen Objekterfassung, basierend auf der Auswertung der Bestimmungselemente eines Dreiecks, nämlich Basislänge und Richtungswinkel (Abschnitt 6.2). Die Interferometrie

verwendet Lichtinterferenzen zur Messung von sehr geringen Distanzunterschieden bis  $1/100$  der Wellenlänge (Schwarte u. a. 1999). Interferometrische Verfahren sind z. B. die Speckle-Interferometrie zur Vermessung kleinster Objektverformungen (z. B. Purde 2006) oder die interferometrische Längenmessung als das genaueste Längenmessverfahren mit der höchsten Auflösung (z. B. Möser u. a. 2000). Sie zeichnen sich in der Regel durch einen sehr hohen Messaufwand aus, seien an dieser Stelle der Vollständigkeit halber genannt und im Folgenden nicht weiter betrachtet. Die Verfahren der Laufzeitmessung messen die Laufzeit eines Signals direkt oder indirekt durch gepulste bzw. optische Modulationstechniken (Abschnitt 6.3).

Für ein alternatives Klassifikationsschema von Verfahren der 3D-Objekterfassung sei auf (Maas 1992b) verwiesen.

## 6.2 Triangulation

Die Triangulation ist das am weitesten verbreitete Verfahren zur optischen dreidimensionalen Form- erfassung und steht in direkter Konkurrenz zu der im Rahmen dieser Arbeit eingesetzten distanz- messenden Kamera. Triangulationsverfahren lassen sich nach [Abbildung 6.2](#) primär unterscheiden in aktive Triangulationstechniken wie das Lichtschnittverfahren ([Abschnitt 6.2.1](#)) und passive Tri- angulation durch Techniken der Stereo-Photogrammetrie ([Abschnitt 6.2.2](#)). In der Computergrafik werden auch häufig Schattierungs- und Silhouettenverfahren eingesetzt ([Abschnitte 6.2.3](#) und [6.2.4](#)). Darüber hinaus können Triangulationsverfahren noch in Fokussierungstechniken und Theodolitmess- verfahren differenziert werden, welche an dieser Stelle nur der Vollständigkeit halber genannt sind.

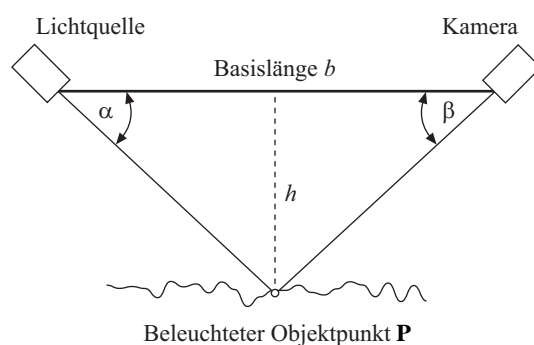


**Abb. 6.2:** Triangulationsverfahren (nach Schwarte u. a. 1999).

### 6.2.1 Aktive Triangulationsverfahren

Im einfachsten Fall der Triangulation wird ein Lichtpunkt mit z. B. einem Laser auf eine Objekt- oberfläche projiziert und nach der Reflexion am Objekt auf einem positionsempfindlichen Detektor registriert, z. B. auf dem Sensorelement einer Kamera ([Abbildung 6.3](#)). Die Länge  $b$  der Triangu- lationsbasis und der Winkel  $\alpha$  zwischen der Basis und dem Sender sind fest und durch ein geeig- netes Kalibrierungsverfahren bekannt. Die Größe des fehlenden Winkels  $\beta$  ist direkt abhängig von der

Entfernung  $h$  des Objektpunkts  $\mathbf{P}$  zur Basis  $b$ . Das Triangulationsdreieck Sender-Objekt-Empfänger ist somit eindeutig definiert und die Berechnung der 3D-Koordinaten des Objektpunkts  $\mathbf{P}$  möglich. Ein limitierender Faktor bei der erreichbaren Objektpunktgenauigkeit (in Aufnahme­richtung) ist das Basis-Höhen-Verhältnis  $h : b$ , welches die Schnittgeometrie bei der Punktbestimmung beschreibt. Schleifende Strahlenschnitte ergeben sich dann, wenn die Basis im Verhältnis zum Abstand klein ist. Somit leiden alle aktiven Triangulationsverfahren unter dem notwendigen räumlichen Abstand zwischen Lichtquelle, Kamera und Objektoberfläche. Gleiches gilt selbstverständlich auch für die im Weiteren diskutierten passiven Triangulationsmethoden unter der Verwendung einer oder mehrere Kameras anstelle einer aktiven Lichtquelle.

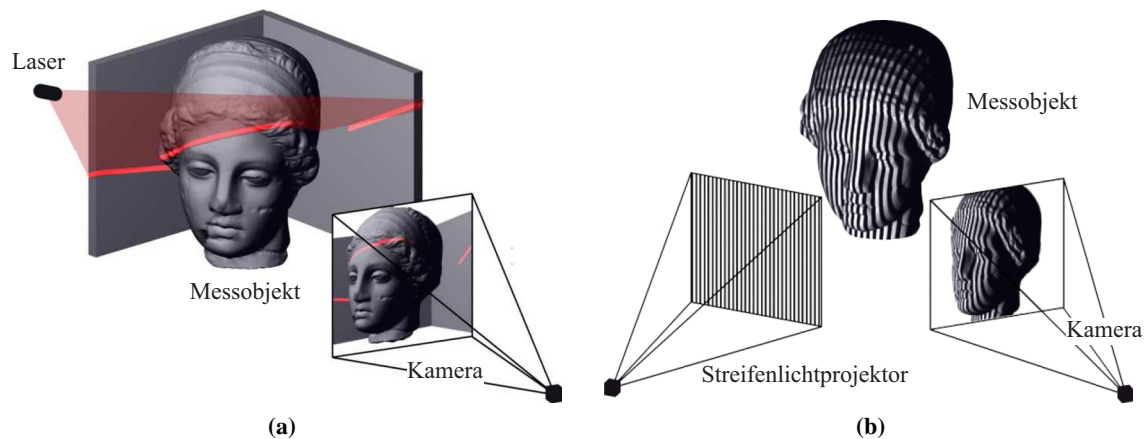


**Abb. 6.3:** Prinzip der aktiven optischen Triangulation.

Beim Lichtschnittverfahren wird der Laserstrahl durch eine Zylinderlinse zu einer Laserlinie aufgeweitet (Abbildung 6.4a). Aus der Blickrichtung des Empfängers erscheint diese deformiert auf dem Messobjekt. Die Abweichung von der Geradheit ist ein Maß für die Tiefenausdehnung. Aus einer Bildaufnahme lassen sich so die 3D-Koordinaten aller Objektpunkte berechnen, die auf der aufgespannten Lichtebene liegen. Dieser Übergang von einer ein- auf eine zweidimensionale Lichtprojektion bedeutet einen erheblichen Zeitgewinn für die Digitalisierung. Zur vollständigen Erfassung wird nun entweder das Objekt oder der Lichtschnittsensor bewegt. Die Positionierungsgenauigkeit muss dabei ausreichend hoch sein, da alle Fehler bei der Bestimmung der absoluten Lage der Triangulationsdreiecke direkt in die Messung eingehen.

Die logische Erweiterung der vorgestellten Ansätze ist die dreidimensionale strukturierte Beleuchtung des gesamten Objekts. Ein Projektor stellt hierbei die Lichtquelle dar und projiziert aktiv ein z. B. streifen- oder sinusförmiges bzw. zur Auflösung von Mehrdeutigkeiten ein grau-kodiertes Streifenmuster auf die Objektoberfläche (Abbildung 6.4b). Die Geometrie des projizierten Musters und des Videoprojektors als inverse Kamera sind durch eine Systemkalibrierung bekannt. Die Objektoberfläche kann nun effektiv mit bereits einer Kamera digitalisiert und anschließend durch Phasen- oder Parallaxenmessungen im detektierten Muster rekonstruiert werden. Der Einsatz von Mehrkamerasystemen mit aktiver Musterprojektion erhöht die Zuverlässigkeit und Robustheit der Messung und ermöglicht die Erfassung bewegter Objekte.

Der Einsatz von optischen Triangulationsverfahren erfolgt bspw. über Förder- oder Fließbändern



**Abb. 6.4:** Prinzipien des Lichtschnittverfahrens (a) und der strukturierten Beleuchtung (b) (nach Winkelbach 2006).

in der industriellen Qualitätskontrolle (Mollath & Fatehi 1994), zur dynamischen Erfassung von Wasserspiegelhöhen im wasserbaulichen Versuchswesen (Mulsow u. a. 2008) oder zur 3D-Modellierung anatomischer Strukturen von Pflanzen, Tieren und Menschen zu Zwecken der Computeranimation (Breiner 2005).

### 6.2.2 Passive Triangulationsverfahren

Passive Triangulationsverfahren umfassen die Aufnahme- und Auswerteverfahren der Photogrammetrie. Bei bekannter geometrischer Objektform können 3D-Koordinaten bereits aus einem Bild berechnet werden (Einbildauswertung; Abschnitt 5.3). Liegen keine Objektinformationen vor, werden zur 3D-Rekonstruktion eines statischen Objekts mit einer Kamera zwei Bilder aus konvergenten Blickrichtungen aufgenommen. Die Aufgaben der klassischen Zweibildauswertung sind die Identifikation und Messung diskreter Punkte in den Messbildern durch Verfahren der Bildzuordnung. Solche Zuordnungsverfahren sind bspw. merkmalsgestützte Interest-Operatoren oder flächenbasierte Ansätze wie die Kleinste-Quadrate-Anpassung (Abschnitt 7.2). Die Berechnung der 3D-Koordinaten erfolgt durch räumliches Vorwärtseinschneiden auf der Basis des zugrunde liegenden Transformationsmodells, z. B. der Kollinearitätsgleichungen (5.7) bei einer Zentralprojektion (Abschnitt 5.2). In Abhängigkeit von der Anzahl der gemessenen Bildpunkte (und der räumlichen Sensorauflösung) kann die Objektoberfläche prinzipiell sehr genau approximiert werden. Das Verfahren der Stereophotogrammetrie kann durch die Aufnahme von im Prinzip beliebig vielen Bildern zu einer Bündeltriangulation erweitert werden (Mehrbildphotogrammetrie; Abschnitt 5.4).

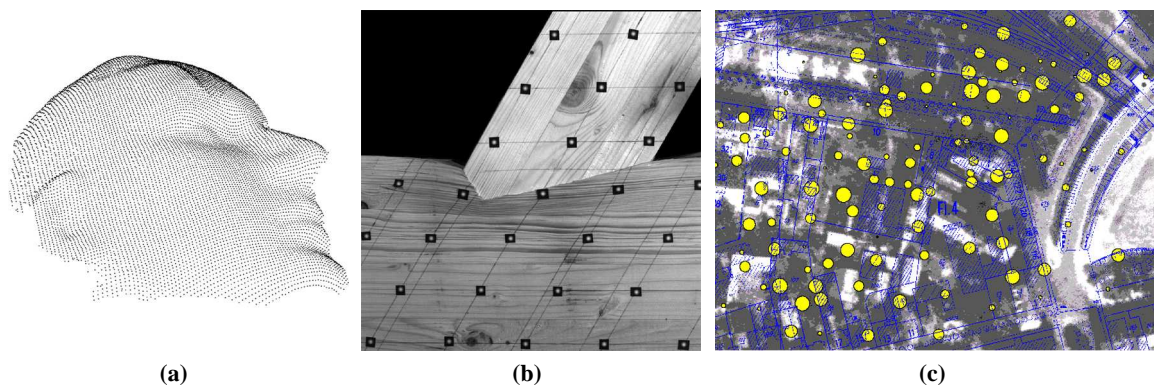
Sollen dynamische Objektzustände erfasst werden, sind Bildsequenzen von zwei oder mehr synchronisierten Kameras auszuwerten. Neben der räumlichen Zuordnung zeitgleich aufgenommener Bilder, ergibt sich auch die Notwendigkeit der Zuordnung auf der Zeitachse. Je nach Aufgabenstellung kann die Lösung der spatiotemporalen Zuordnung hohe Komplexität aufweisen und nicht immer



zu fehlerfreien Lösungen führen (Abschnitt 7.3).

Objekte werden bei einer Größe von etwa 1 m mit Genauigkeiten von unter  $1/10$  mm und bei einer Größe bis ca. 200 m mit Genauigkeiten im cm-Bereich erfasst (Luhmann 2003).

Die Anwendungsmöglichkeiten sind zahlreich und reichen von der Generierung von Oberflächenmodellen menschlicher Körperteile aus durch Mehrbildtechniken gerechneten 3D-Punktwolken (D'Apuzzo 2003; Abbildung 6.5a) über die Detektion und Messung kleinster Verformungen in der Materialwissenschaft (Hampel 2008; Abbildung 6.5b) bis hin zur Herstellung von verzerrungs- und maßstabgetreuen Orthofotos, gerechnet aus digitalen Geländemodellen, die anschließend bspw. mit Sekundärinformationen überlagert werden können (Buhmann u. a. 1996; Abbildung 6.5c). Die Auswertung mon- oder multiokularer Bildsequenzen für Anwendungen der menschlichen 3D-Bewegungsanalyse wird darüber hinaus speziell im Kapitel 12 behandelt.



**Abb. 6.5:** Anwendungen der photogrammetrischen Bildzuordnung: (a) 3D-Punktwolke eines Gesichts (D'Apuzzo 2003). (b) Verformungsmessung bei Belastungsversuchen an Holzversatzkonstruktionen (Hampel 2008). (c) Ein mit Katasterdaten überlagertes Orthofoto (Buhmann u. a. 1996).

### 6.2.3 Schattierungsverfahren

Schattierungsverfahren (*Shading*) kommen überwiegend in der Computergrafik zu Zwecken der Oberflächensimulation zum Einsatz. Der Grauwert ergibt sich dabei aus der Oberflächenneigung relativ zu einer frei gewählten Beleuchtungsrichtung. Umgekehrt können auf den Reflexionseigenschaften des Objekts basierende Verfahren aber auch zur 3D-Oberflächenrekonstruktion angewendet werden, die Form einer Oberfläche also auch aus der lokalen Orientierung der Oberflächenelemente bestimmt werden. Horn beschrieb dieses Prinzip bereits in seiner 1970 erschienenen Dissertationsschrift und bezeichnete den Ansatz als *Shape from Shading* (SfS).

Die Orientierung eines kleinen Oberflächenelements ist gegeben über den darauf senkrecht stehenden Normalenvektor (Abbildung 6.6a). Die Oberflächennormale ist aus der Albedo im Bild bestimmbar, also aus dem Maß des Rückstrahlvermögens der diffus reflektierenden Objektfläche. Dabei wird ausgenutzt, dass die Albedo vom Winkel zwischen der Oberflächennormalen und der Richtung der Beleuchtung abhängt. Unter der Annahme einer konstanten und ortsunabhängigen Beleuchtung,

einer kontinuierlich verlaufenden Oberflächenneigung sowie einer bekannten Reflexion (z. B. Lambertsches Reflektionsmodell) können bereits aus einem Einzelbild die 2-komponentigen Oberflächennormalen für jedes Pixel ermittelt, Linienscharen maximaler Grauwertänderungen zusammengefasst in Nadeldiagrammen analysiert und über Integrationstechniken zu einem vollständigen Oberflächenmodell verknüpft werden (Abbildung 6.6b). Ausführliche Rechenvorschriften werden in (Klette u. a. 1999) gegeben. Ein großer Vorteil dieser Vorgehensweise ist der Wegfall des Korrespondenzproblems (Abschnitt 7.3). Es können also Oberflächen aus texturlosen Regionen modelliert werden. Andererseits schränken die zu treffenden und meist höchst restriktiven Annahmen über z. B. vorherrschende Beleuchtungs- und Reflexionsmodelle, Oberflächenstetigkeiten oder die Grundform der Oberfläche selbst SfS als universell einsetzbares Verfahren zur 3D-Rekonstruktion stark in Genauigkeit und Zuverlässigkeit ein (Maas 1997).

Die Erweiterung um zusätzliche Eingabebilder heißt *Photometric Stereo*<sup>3</sup> (Woodham 1978) und behebt die Unterbestimmtheit durch die Verwendung multipler Lichtquellenpositionen. Die Reflexionsabbildungen von aus verschiedenen Positionen punktförmig strahlenden Lichtquellen werden sequentiell von einem Standpunkt aufgenommen.

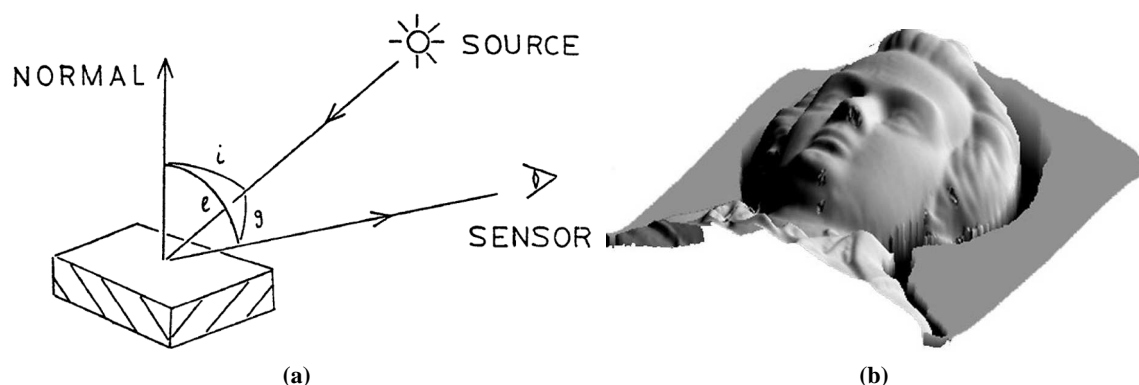
Alternativ können zur Erfassung bewegter Objektoberflächen mit einer Kamera synchronisierte alternierende Lichtquellen (Piironen u. a. 2000) oder auch drei Lichtquellen unterschiedlicher Farbe und eine 3-Chip Farbkamera verwendet werden (*Colour Photometric Stereo*; M. L. Smith & Smith 2005). Unter bekannter Orientierung der Lichtquellen und der Kamera kann die Oberfläche vollständig und ohne zusätzliche Annahmen über bspw. Objektform und -symmetrie durch den Schnitt von Kurvenscharen gleicher Grauwerte und somit auch gleicher Oberflächenneigungen aus mehreren Bildern rekonstruiert werden. Prinzipbedingt ergeben sich Einschränkungen im globalen Genauigkeitspotential (Maas 1997). Nachteilig ist weiterhin der große Messaufwand, bedingt durch eine strenge radiometrische und geometrische Kalibrierung. Der erhebliche Rechenaufwand sowohl für die Bestimmung der Schnittpunkte als auch für die Berechnung der Iso-Intensitätskonturen komplexer Reflexionsmodelle ist bei einigen Anwendungen problematisch.

Li (1992) rekonstruiert aus monokularen Sonarbildern die Oberfläche des Meeresgrunds, Pradhan u. a. (2010) aus Satellitenbildern die Oberfläche der Erde. In (W. A. Smith & Hancock 2010) werden die Reflexionseigenschaften von Gesichtern in Verbindung mit der Bestimmung der gesamten Gesichtsoberfläche ermittelt. Mitchell (2010) untersucht im Rahmen einer Machbarkeitsstudie die Eignung von SfS im Kontext zahnmedizinischer Vermessungsarbeiten.

Der photometrische Ansatz kommt also in den unterschiedlichsten Bereichen zur Anwendung, oft auch als komplementäres Verfahren. Eine Auswahl von in Kombination mit der Tiefenbildtechnologie durchgeführten Arbeiten wird im Kapitel 13 vorgestellt.

---

<sup>3</sup> Es sei angemerkt, dass diese erweiterte Methode streng genommen der Gruppe der aktiven Triangulationsverfahren zuzuordnen ist, da Licht in die Szene geworfen wird. Aufgrund ihrer Nähe zu konventionellen SfS-Techniken wird sie aber an dieser Stelle nicht beschrieben.



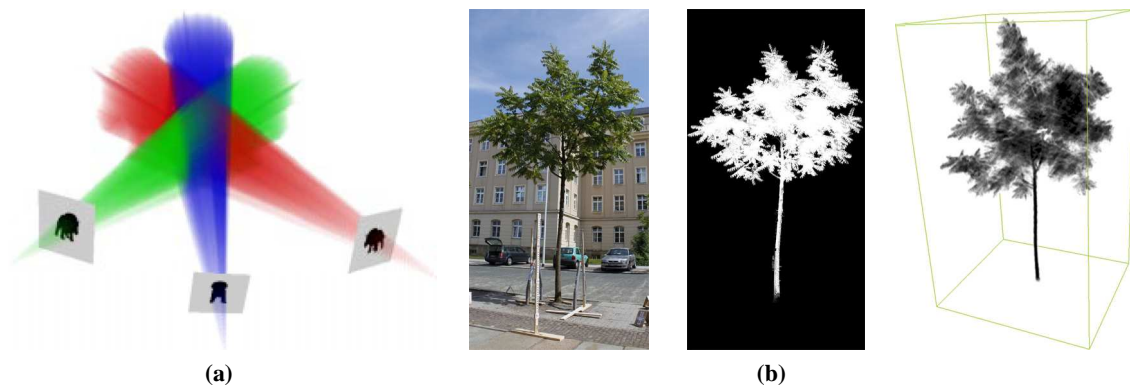
**Abb. 6.6:** (a) Die Orientierung eines Oberflächenelements ist mathematisch definiert über deren Oberflächennormale, also den drei photometrischen Winkeln  $i$ ,  $e$  und  $g$  (Horn 1970). (b) Texturierte Mozartbüste als Ergebnis von *Photometric Stereo* (Klette u. a. 1999).

#### 6.2.4 Silhouettenschnittverfahren

Eine volumetrische Rekonstruktion aus den Umrissen eines Objekts wird als *Shape from Silhouette* bezeichnet (Laurentini 1994). Bei dem Verfahren werden im einfachsten Fall Objektaufnahmen aus verschiedenen Richtungen durch Methoden der Bildverarbeitung binär in Vorder- und Hintergrund segmentiert. Jedes Vordergrundpixel jeder Kameraaufnahme repräsentiert die Silhouette des zu rekonstruierenden Objekts und wird über die bekannten Orientierungsparameter in den Objektraum projiziert (Abbildung 6.7a). Der Schnitt aller Projektionsstrahlen (Silhouettenschnitt) erzeugt die dreidimensionale visuelle Hülle des abgebildeten Objekts und grenzt dessen Volumen im Raum ein. Die Voxel der Hülle werden anschließend iterativ entsprechend ihrer Binärwerte im Bild entweder transparent (Hintergrund) oder opak (Vordergrund) dargestellt.

Auf dem Silhouettenschnittverfahren basierende Ansätze sind aufgrund ihrer Schnelligkeit und Robustheit vor allem in der Computergrafik weit verbreitet. Die Repräsentationsqualität hängt neben der Segmentierungsgenauigkeit primär von der Standpunktanzahl und der Komplexität des Objekts ab. Laurentini (1997) beschreibt dies detailliert. Einschränkungen ergeben sich bei der Rekonstruktion konkaver Objektstellen, welche u. U. über die Verwendung von Farbinformationen kompensiert werden können (Kutulakos & Seitz 2000).

Ein Beispiel für die fotorealistische volumetrische Rekonstruktion eines Baumes wird in Abbildung 6.7b gezeigt. Vock u. a. (2010) segmentieren hier zunächst die Aufnahmen eines Mehrbildverbands nicht binär, sondern in Grautonskalen entsprechend ihrer Opazität (*Alpha Matting*). Das Baumvolumen wird durch eine Vorwärtsprojektion der einzelnen Voxel in die Bildebenen mit Opazitätswerten initialisiert und anschließend durch iterative Rückprojektionen verbessert (*Ray Casting*). In der *Computer Vision* ebenfalls weit verbreitet ist der Einsatz von Silhouettenschnittverfahren bei der Separation und anschließenden Verfolgung sich bewegender Objekte oder Personen. Beispiele hierfür folgen im Kapitel 13.



**Abb. 6.7:** (a) Prinzip des Silhouettenschnittverfahrens (Buehler u. a. 1999): Die Schnittpunkte der drei Konen definieren die visuelle Hülle des Objekts. (b) Volumetrische Rekonstruktion eines Baums (Vock u. a. 2010): Aufnahme eines Mehrbildverbands (links). Das entsprechend nach Vorder- und Hintergrundzugehörigkeit in eine Grautonskala segmentierte Eingabebild (Mitte). Beispiel eines rekonstruierten Baums (rechts).

## 6.3 Laufzeitmessung

Die Laufzeitmessung (*Time-of-Flight, ToF*) auf der Basis von LIDAR-Systemen (*Light Detection and Ranging*) unterscheidet sich nach zwei physikalischen Prinzipien (Joeckel & Stober 1999; Abbildung 6.8): Beim Impulsmessverfahren (Abschnitt 6.3.1) wird die Distanz direkt, beim Phasenvergleichsverfahren (Abschnitt 6.3.2) indirekt über die Messung der Laufzeit eines ausgesendeten Signals ermittelt. In Abhängigkeit von der Strahlungsenergie können Objekte in einer Entfernung von wenigen Metern bis hin zu  $2 \cdot 10^4$  km bei GPS-Satelliten (Globales Positionierungssystem) vermessen werden.

Neben z. B. elektronischen Tachymetern ist der terrestrische Laserscanner ein wichtiger Vertreter, dessen Messprinzip auf der elektrooptischen Laufzeitmessung basiert. Als weit verbreitetes 3D-Messsystem zur großflächigen Erfassung von z. B. Gebäuden steht auch er in Konkurrenz zum monosensoriellen 3D-Kamerasystem und wird im Abschnitt 6.3.3 kurz vorgestellt.

### 6.3.1 Impulsmessverfahren

Beim Impulsmessverfahren (Pulsmodulation) wird die Laufzeit eines ausgesendeten, am Objekt reflektierten und über einen Fotodetektor empfangenen Lichtimpulses gemessen (Abbildung 6.9). Die gesuchte Distanz  $D$  berechnet sich direkt durch Multiplikation der halben Laufzeit  $t$  mit der Ausbreitungsgeschwindigkeit  $c$  von Lichtwellen in Luft.

Die direkte Laufzeitmessung muss für das Erreichen von geodätisch interessanten Distanzmessegenauigkeiten sehr präzise durch kostenintensive Uhren erfolgen. Für z. B. eine Genauigkeit von 5 mm hat sie  $0,33e^{-10}$  s zu betragen.

### 6.3.2 Phasenvergleichsverfahren

Beim Phasenvergleichsverfahren wird die Laufzeit nicht direkt bestimmt, sondern über eine Messung der Phasenverschiebung zwischen der Modulation einer aktiven Beleuchtung und der Modulation des reflektierten Empfängersignals abgeleitet (Abbildung 6.10). Ein Sender strahlt kontinuierlich eine Trägerwelle (*Continuous Wave*) mit einem periodisch aufmodulierten Messsignal aus. Die aufmodulierte und meist sinusförmige Welle dient dabei als Maßeinheit. Die zugehörige Maßstabwellenlänge  $\lambda$  kann über das Verhältnis zwischen der Signalgeschwindigkeit  $c$  und einer vorgegebenen Modulationsfrequenz  $f$  berechnet werden:

$$\lambda = \frac{c}{f} = \frac{c_0}{n \cdot f} \quad (6.1)$$

mit

$c_0$  Lichtgeschwindigkeit im Vakuum

$n$  Brechzahl der Atmosphäre

Die Distanzmesser bestehen aus einem Sender, einem Empfänger und einem Phasenmesser. Die vom Sender erzeugte und ausgesendete Messwelle  $y_A$  wird vom Objekt zurückreflektiert. In Abhängigkeit von der durchlaufenen doppelten Distanz  $2D$  erfährt das empfangene Signal  $y_E$  eine Phasenverschiebung  $\varphi$ :

$$\varphi_{2D} = N \cdot 2\pi + \Delta\varphi \quad (6.2)$$

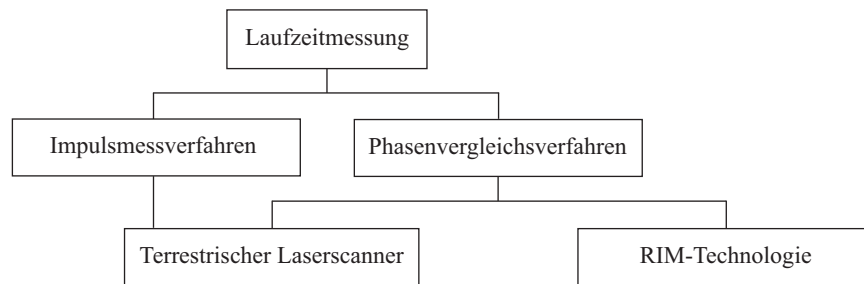
In der Gleichung (6.2) entspricht  $2\pi$  einer vollen Wellenlänge  $\lambda$  und das Vielfache  $N$  ist die Anzahl der vollen Wellenzüge. Die Phasendifferenz  $\Delta\varphi$  vom Ende des letzten vollen Wellenzugs bis zum Empfänger wird vom Phasenmesser durch die Verschiebung des empfangenen Signals  $y_E$  gegen ein zuvor vom ausgesendeten Signal  $y_A$  abgetrenntes Referenzsignal bestimmt. Die durchlaufene Strecke  $2D$  setzt sich somit zusammen zu

$$2D = N \cdot \frac{\lambda}{2} + \frac{\Delta\lambda}{2} \quad (6.3)$$

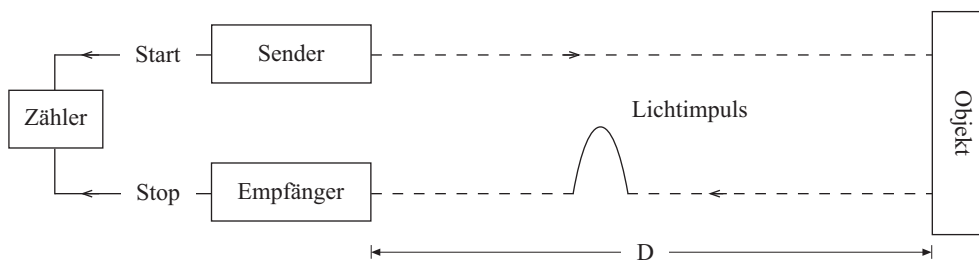
mit

$$\Delta\lambda = \frac{\Delta\varphi}{2\pi} \cdot \lambda$$

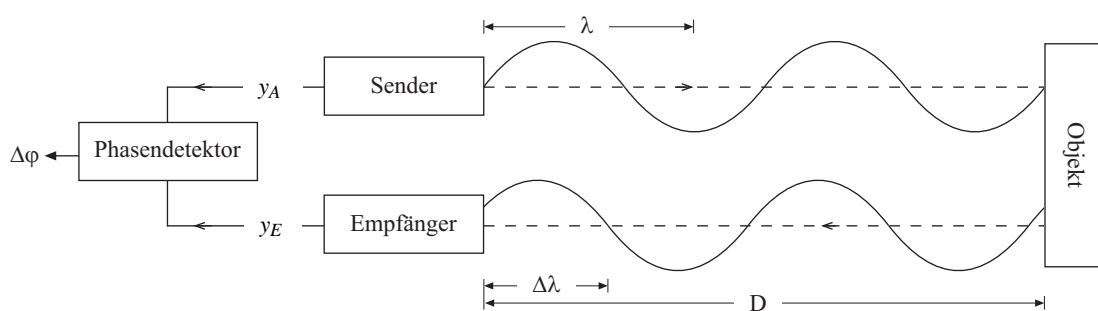
$D$  ist eindeutig bestimmbar und genau dann gleich dem Wellenreststück  $\Delta\lambda$ , wenn es kleiner oder gleich der halben Wellenlänge  $\lambda$  ist. Sollen größere Distanzen gemessen werden, sind mindestens zwei Messungen mit verschiedenen Maßstabwellenlängen notwendig. Zur weiteren Vertiefung der Thematik sei auf (Joeckel & Stober 1999) verwiesen.



**Abb. 6.8:** Verfahren zur Laufzeitmessung.



**Abb. 6.9:** Messprinzip des Impulsmessverfahrens (nach Joeckel & Stober 1999).



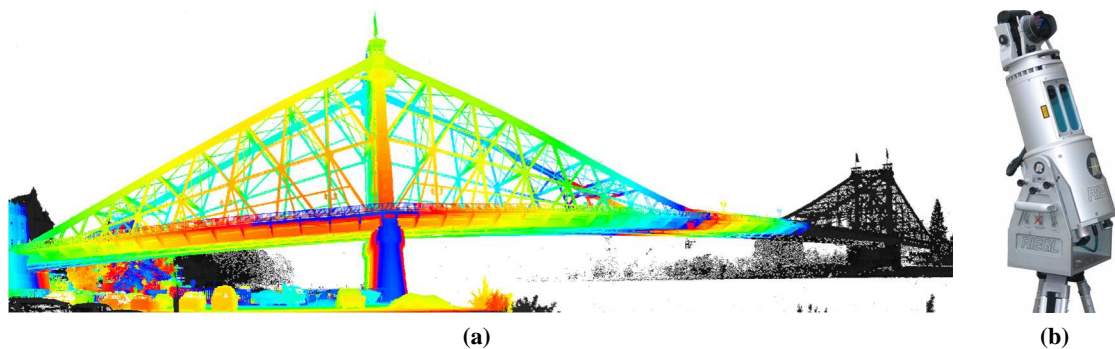
**Abb. 6.10:** Messprinzip des Phasenvergleichsverfahrens (nach Joeckel & Stober 1999).



### 6.3.3 Laserscanning

Beim Laserscanning wird die Oberfläche eines Objekts durch sequentielle Abtastung dreidimensional erfasst. Es existieren Instrumente, die auf dem Impulsmessverfahren, dem Phasenvergleichsverfahren oder einer Kombination beider Prinzipien beruhen. Wenige Laserscanner setzen auch das Triangulationsverfahren ein. Die räumlichen Objektkoordinaten ergeben sich aus der elektrooptischen Distanzmessung sowie den zusätzlich erfassten Horizontal- und Vertikalwinkeln durch polares Anhängen und liegen in einem Winkelraster vor (Kahmen 1997). Die 3D-Punkte können zusätzlich mit aus der Distanzmessung resultierenden monochromatischen Intensitätsinformationen überlagert werden. Eine ausführliche Darstellung des Funktionsprinzips und weiterer Aspekte der Laserscanning-Technologie finden sich in (Vosselman & Maas 2010).

Laserscanner werden in Flugzeugen für z. B. die Erstellung von digitalen Geländemodellen (Maas 2005a) und auf der Erdoberfläche zur Vermessung von Objekten im Nahbereich (Scheller & Schneider 2006; Abbildung 6.11a) eingesetzt. Terrestrische Laserscanner (TLS; Abbildung 6.11b) zeichnen sich durch eine schnelle 3D-Erfassung großer Objektoberflächen mit einer hohen Punktdichte über Entfernungen von bis zu 6.000m aus. Schneider (2009) stellt jedoch fest, dass die Aufnahmezeiten pro Standpunkt, in Abhängigkeit vom gewählten Punktabstand, dem Messprinzip des TLS und der Objektgröße, einige Minuten betragen können. Die Genauigkeit einer in 10m Entfernung gemessenen 3D-Koordinate beträgt für eine typische TLS-Systemkonfiguration ca. 3mm bis 10mm (Luhmann 2003).



**Abb. 6.11:** 3D-Laserscanning: (a) Entfernungskodierte 3D-Punktwolke des Blauen Wunders in Dresden. (b) Terrestrischer Laserscanner LMS-Z620, RIEGL Laser Measurement Systems GmbH (Rieg 2010).

## 6.4 Diskussion

Das Kapitel bot einen kurzen Überblick über die in der optischen Messtechnik verbreiteten Messverfahren und deren Anwendungsmöglichkeiten. Die vorgestellten Verfahren arbeiten alle berührungslos und basieren auf der Erfassung von aktiv ausgesendeter oder durch eine weit entfernte Lichtquelle

passiv emittierter Strahlung, welche vom Objekt teilweise zurück reflektiert und bildgebend erfasst wird. Die anschließend zur Rekonstruktion von räumlichen Objekten genutzten Messgrößen sind bei photometrischen Verfahren die Grauwerte selbst, bei allen anderen Triangulationsansätzen sind es die Bildkoordinaten oder binäre Streifen- bzw. Phaseninformationen. Bei den Verfahren der Laufzeitmessung werden die Distanz- und Winkelinformationen verwendet.

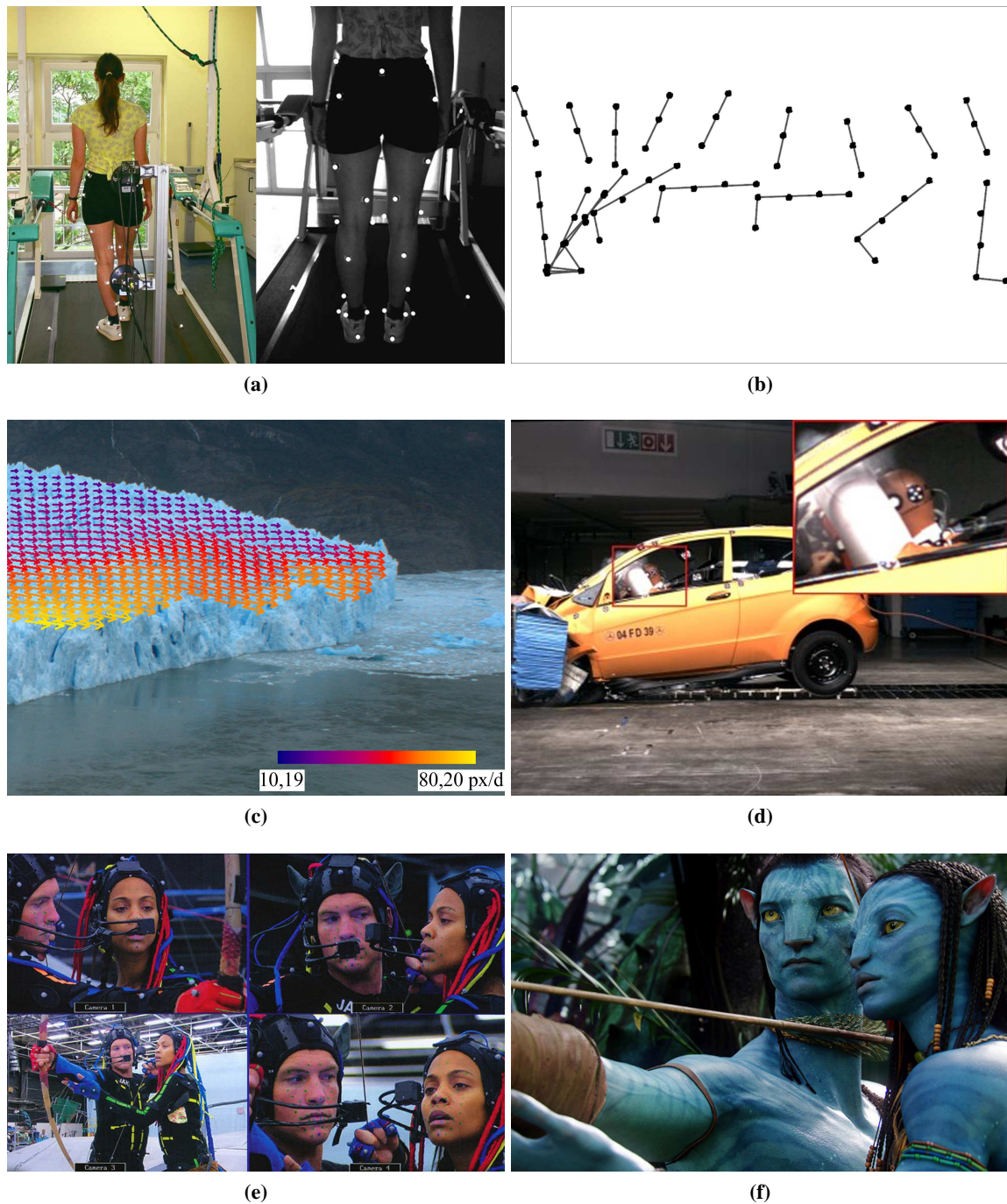
Die unterschiedlichen Prinzipien, Messanordnungen und Genauigkeitspotentiale wurden erläutert sowie die spezifischen Vor- und Nachteile der Techniken diskutiert. Ausführlicher behandelt wurden die Grundlagen für das Verständnis des Phasenvergleichsverfahrens, auf welchem das Messprinzip von der im Rahmen der vorliegenden Arbeit zur spatialen und temporalen Oberflächenerfassung verwendeten und im [Kapitel 8](#) detailliert vorgestellten 3D-Kamera beruht. Relevante Unterschiede ergeben sich prinzipbedingt durch den technischen Aufwand, die Simultanität des Messvorgangs, die erreichbare spatiale und/oder temporale Auflösung sowie die Messgenauigkeit und -zuverlässigkeit. An dieser Stelle sollen sie jedoch nicht untereinander, sondern besonders bezüglich der Auswertung dynamischer und speziell menschlicher Bewegungsvorgänge im [Abschnitt 8.5](#) direkt mit einer 3D-Kamera verglichen werden.



## Bewegungsanalyse

Die Bewegung ist eines der wichtigsten Merkmale von Bild- bzw. Videosequenzen und kann zur Identifizierung und Charakterisierung von Objekten sowie zum Bildverstehen (*Image Understanding*) verwendet werden. Die Bildsequenzanalyse offenbart dynamische Prozesse, also die Beziehungen zwischen räumlichen Bildmerkmalen und zeitlichen Veränderungen, und ist für eine Vielzahl von Anwendungen unabdingbar. In der Rehabilitationsmedizin werden Fortschritte bspw. bei der Wiederherstellung der Funktionsfähigkeit des Bewegungsapparats über Ganganalysen gemessen ([Abbildung 7.1a](#) und [7.1b](#)). In der Glaziologie kommen Verfahren der Bildsequenzanalyse zum Einsatz, um die Fließgeschwindigkeiten von Gletschern zu überwachen ([Abbildung 7.1c](#)); in der Automobilindustrie, um das Unfallverhalten von Kraftfahrzeugen (Crashtest) aufzunehmen und analysieren zu können ([Abbildung 7.1d](#)); in der Unterhaltungsindustrie zur Erfassung und digitalen Weiterverarbeitung menschlicher Bewegungen (*Motion Capturing*; [Abbildung 7.1e](#) und [7.1f](#)).

Im [Abschnitt 7.1](#) sollen zunächst die Grundlagen aus dem Bereich Bewegung sowie deren Detektion und Modellierung bereitgestellt werden. Der folgende [Abschnitt 7.2](#) erläutert die Prinzipien einer Bildzuordnung und kategorisiert die Verfahren aus Sicht der *Computer Vision* und der Photogrammetrie. Die Betrachtung von den Standpunkten unterschiedlicher Wissenschaftsgebiete erscheint dem Autor sinnvoll und notwendig, um die bestehende, aber in der Literatur zumeist nicht thematisierte Verbindung zwischen dem Begriff *Optischer Fluss* in der *Computer Vision* und der *Bildzuordnung* als Ausdruck für die photogrammetrische Messung homologer Punkte in räumlich oder zeitlich verknüpften Bildern herzustellen. Nach der Behandlung grundlegender Probleme, die bei der Zuordnung korrespondierender Bildpunkte auftreten können ([Abschnitt 7.3](#)), werden in den abschließenden [Abschnitten 7.4](#) bis [7.6](#) drei wichtige Verfahren zur Berechnung von Bewegungsvektorfeldern erläutert: Die Kreuzkorrelation als ein sehr weit verbreitetes und einfach zu implementierendes Bildzuordnungsverfahren wird im [Abschnitt 7.4](#) kurz vorgestellt. Der [Abschnitt 7.5](#) behandelt die lokale und die globale Methode des Optischen Flusses, welche als Standard zur Analyse von Bewegungen im Bereich der Computergrafik gilt und in erweiterter Form auch zur Analyse von 3D-Kameraaufnahmen genutzt wird ([Abschnitt 13.3](#)). Abschließend werden das *2D Least Squares Matching* (LSM) und darauf basierende wichtige Erweiterungen im [Abschnitt 7.6](#) diskutiert. 2D-LSM gilt als eines der bedeutendsten Bildzuordnungsverfahren der digitalen Photogrammetrie und wird im Rahmen der vorliegenden Arbeit zur Auswertung von 3D-Kamerabildsequenzen adaptiert und weiterentwickelt ([Kapitel 14](#)).

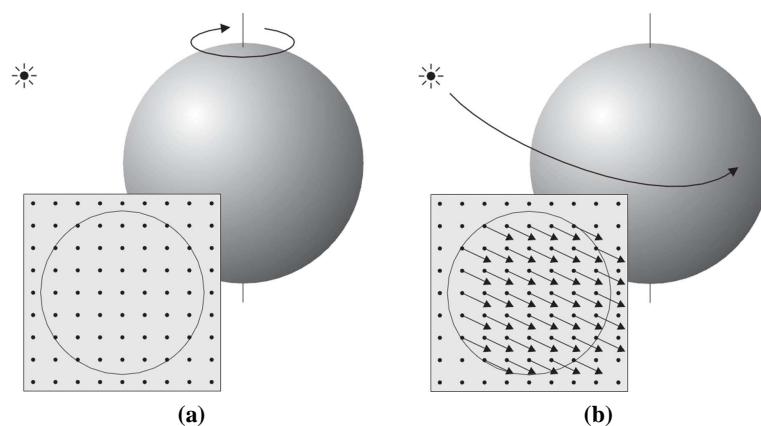


**Abb. 7.1:** Verfahren der automatischen Bewegungsanalyse finden in unterschiedlichen Disziplinen Anwendung: (a) Stereokamerasystem zur Bewegungsanalyse am Laufband (Hofmann 2008). (b) Visualisierung des rekonstruierten Gangzyklus im Inertialsystem des Laufbandes (Putze 2008). (c) Bewegungsvektorfeld einer Gletscheroberfläche (Maas u. a. 2010). (d) Positionsbestimmung eines Dummies beim Crashversuch (Gall u. a. 2008). (e, f) Für den in größtenteils virtuellen Studios gedrehten Film *Avatar – Aufbruch nach Pandora* (2009) wurden durch *Motion-Capture*-Verfahren Schauspieler fotorealistisch erfasst und anschließend computeranimiert (Fanpop 2011).

## 7.1 Übersicht

### 7.1.1 Bewegung

Bewegungen werden naturgemäß mit Veränderungen verbunden. In der Bildanalyse treten solche Veränderungen in Form von räumlichen Grauwertänderungen auf. Eine stattgefundenene Bewegung kann folglich auch nur an den Stellen im Bild detektiert und im Anschluss numerisch erfasst werden, an denen sich die Grauwerte zeitlich ändern. Der Umkehrschluss ist in der Praxis jedoch nicht gegeben. Bewegungen können zwar Bildänderungen hervorrufen, allerdings wird nicht jede Bildänderung durch eine Objektbewegung verursacht. Die Beziehungen zwischen der sichtbaren Bewegung im Bild und der realen Bewegung im Objektraum sind dann äquivalent, wenn die Objekte während ihrer Bewegung in der Szene nicht die Beleuchtung in der Bildebene verändern, was lediglich in wenigen Fällen vorkommt. Intensitätsänderungen können zum Beispiel durch eine sich ändernde Lichtquelle hervorgerufen werden. Ebenso möglich ist eine indirekt ausgelöste Variation der Beleuchtungsverhältnisse durch das sich bewegende Objekt. Die [Abbildung 7.2](#) zeigt hierfür zwei Beispiele. Es stellt sich nun die Frage, wie bedeutend die Abweichungen zwischen der realen und der durch den Sensor aufgenommenen Bewegung sind. Für praktische Anwendungen wird meist angenommen, dass sich das tatsächliche Bewegungsfeld nur in geringer Größenordnung von der erfassten Bewegung unterscheidet. Ist dies nicht der Fall oder sind die Genauigkeitsanforderungen der praktischen Auswertung zu hoch, muss die relative Orientierung von Lichtquelle und Sensor zusätzlich berücksichtigt werden.

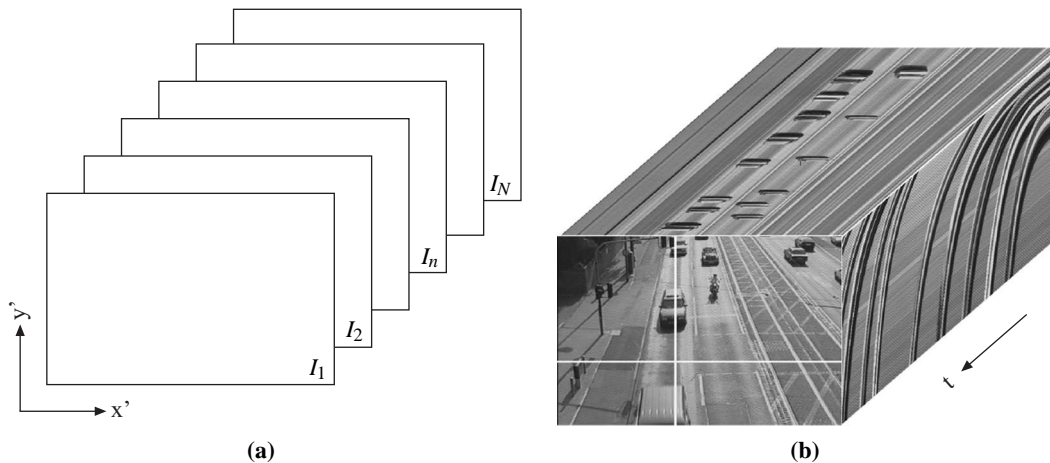


**Abb. 7.2:** Reale vs. visuell wahrnehmbare Objektbewegung (nach [Haußecker & Spies 1999](#)): Es wird eine Kugel mit einer farblich homogenen Oberfläche angenommen. In (a) rotiert die Kugel, ohne dass sich die Intensitäten im Bild ändern. Trotz einer realen Bewegung existiert kein optisches Strömungsfeld, welches durch den bildgebenden Sensor visualisiert werden kann. Im Gegensatz dazu befindet sich die gleiche Kugel in (b) in Ruhe. Die sich bewegende Lichtquelle verursacht Intensitätsänderungen. Diese werden erfasst und bewirken folglich ein optisches Strömungsfeld einer nicht existierenden Objektbewegung.

### 7.1.2 Bildsequenzen

Ein zweidimensionales Einzelbild  $I$  wird durch die diskret auftretenden und durch das Sensorformat beschränkten Pixelkoordinaten  $\mathbf{x}'(x', y')$  beschrieben.  $I(\mathbf{x}')$  kann als Funktion der Intensitätsverteilung des Bilds angesehen werden und liefert das für den Punkt an der Stelle  $\mathbf{x}'$  zugehörige Attribut, in der Regel einen 8bit Grauwert. Die Erweiterung zu einem Bildverband  $I_n(\mathbf{x}')$  mit  $n = [1, N]$  besteht aus  $N$  konsekutiven Einzelbildern  $I_n$ . Bei einem endlichen  $N$  kann nun der Übergang zu einer dreidimensionalen Bildsequenz  $I(\mathbf{x}', t)$  erfolgen (Maas 1997). Neben einem regelmäßigen Ortsraster wird die Zeitkoordinate  $t$  als dritte Dimension eingeführt (Abbildung 7.3). Das Zeitintervall  $\Delta t$  zwischen zwei aufeinander folgenden Bildern wird durch die Bildwiederholrate (Frame Rate) des Sensors und/oder durch das geforderte Diskretisierungsintervall der Anwendung bestimmt. Bildsequenzen in der Form von sequentiellen Mehrbildaufnahmen erlauben somit die Erfassung zeitabhängiger räumlicher Objektbewegungen, wie zum Beispiel Geschwindigkeits- und Beschleunigungsvorgängen.

Zur Aufnahme von Bildsequenzen werden im Allgemeinen digitale (Hochgeschwindigkeits-)Videokameras verwendet (Kapitel 12). Soll eine dreidimensionale Analyse betrieben werden, ist die Erweiterung auf eine mindestens binokulare Aufnahmekonfiguration sowie die Synchronisation und Kalibrierung der verwendeten Kameras notwendig. Alternativ kann auch eine neuartige 3D-Kamera als monosensorielles System zur 3D-Objekterfassung in Videobildrate zum Einsatz kommen (Kapitel 8).



**Abb. 7.3:** Ein Bildblock mit der Zeit in der dritten Dimension als (a) schematische Darstellung und (b) am Beispiel einer Verkehrsszene (nach Jähne 2002).

### 7.1.3 Bewegungsdetektion

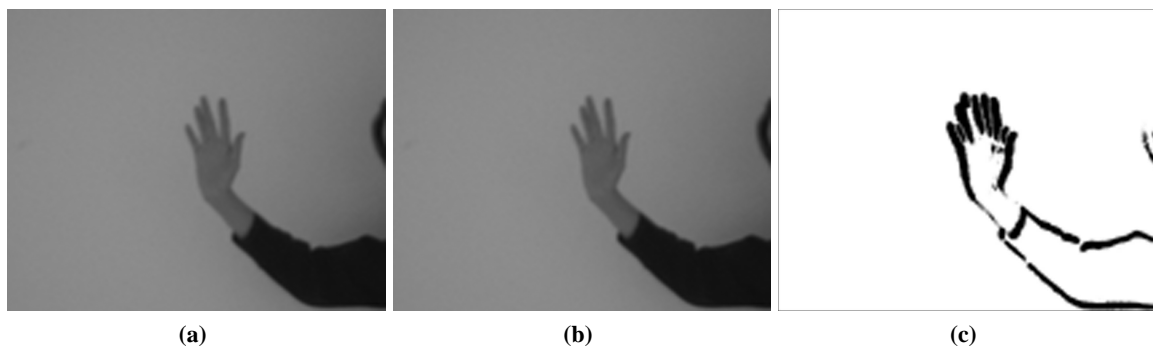
Ein sehr einfaches Verfahren zur Detektion von Bewegungen ist die Erzeugung eines Differenzbilds zweier konsekutiver Bilder einer Bildsequenz. Die Vorschrift zur Erstellung des binären Differenz-

bilds  $\Delta I(\mathbf{x}')$  lautet:

$$\Delta I(\mathbf{x}') = \begin{cases} 255 & \text{für } |I_n(\mathbf{x}', t) - I_{n+1}(\mathbf{x}', t + \Delta t)| \leq \varepsilon \\ 0 & \text{sonst} \end{cases} \quad (7.1)$$

Sobald der Betrag der Grauwertdifferenzen zweier Bilder an der Position  $\mathbf{x}'$  einen zuvor definierten Schwellenwert  $\varepsilon$  überschreitet, wird eine Bewegung detektiert. Es findet also eine durch Bewegung verursachte Grauwertänderung statt. Der sich bewegende Arm in [Abbildung 7.4](#) wird so korrekt, allerdings auch doppelt, registriert. Differenzbilder visualisieren somit Bereiche, in denen eine Bewegung stattfindet. Sie eignen sich allerdings aufgrund der beschriebenen indirekten Bewegungskonzeption nicht für pixel- oder subpixelgenaue Anwendungen und sind lediglich eingeschränkt für zum Beispiel Änderungsdetektionen bei einfachen Überwachungs- und Sicherungsanwendungen brauchbar ([Foresti u. a. 2005](#)).

Zur genauen und zuverlässigen Analyse von Bewegungen in Bildsequenzen eignen sich aus photogrammetrischer Sicht Verfahren der Bildzuordnung ([Abschnitt 7.2](#)), welche auf der Bestimmung von Bewegungsvektorfeldern basieren. Im folgenden [Abschnitt 7.1.4](#) wird die geometrische Modellierung eines solchen Bewegungsvektorfelds aus zwei oder mehr konsekutiven Bildern einer Bildsequenz beschrieben.



**Abb. 7.4:** Zwei konsekutive Bilder (a) und (b) mit ihrem Differenzbild (c).

#### 7.1.4 Modellierung des Bewegungsvektorfelds

Die Projektion einer dreidimensionalen physikalischen Bewegung von sichtbaren Objekten durch die Optik des Sensors auf die Bildebene ordnet jedem Bildpunkt des Bilds  $I_1$  zum Zeitpunkt  $t_1$  einen Verschiebungsvektor  $\Delta \mathbf{x}'(\Delta x', \Delta y')$  zu, der die Bewegungsrichtung und -geschwindigkeit zu seinem korrespondierenden Bildpunkt im Bild  $I_2$  zum Zeitpunkt  $t_2 = t_1 + \Delta t$  angibt. Formal besitzt somit die folgende Beziehung für jedes Pixel an der Position  $\mathbf{x}'$  mit der Intensität  $I(\mathbf{x}', t)$  Gültigkeit:

$$I_1(\mathbf{x}', t_1) = I_2(\mathbf{x}' + \Delta \mathbf{x}', t_1 + \Delta t) \quad (7.2)$$



Die Gesamtheit der Verschiebungsvektoren bildet ein 2D-Verschiebungsvektorfeld  $\xi(\mathbf{x}', t)$ . Dessen zeitliche Ableitung resultiert in einem 2D-Bewegungs- bzw. Geschwindigkeitsvektorfeld  $f(\mathbf{x}', t) = \partial_t \xi(\mathbf{x}', t)$  und wird in der digitalen Bildverarbeitung auch als Optischer Fluss (*Optical Flow*) bezeichnet.

Eine Bewegung wird durch die zeitliche Änderung von Intensitätsmustern in der Bildebene visualisiert. Mathematisch-geometrisch modelliert werden kann dieser Sachverhalt durch eine kontinuierliche Variation der Intensitäten  $I(\mathbf{x}', t)$  in Abhängigkeit von Ort und Zeit. Unter der Voraussetzung kleiner Verschiebungen  $\Delta \mathbf{x}'$  und geringer zeitlicher Unterschiede  $\Delta t$  kann das Verschiebungsvektorfeld  $\xi(\mathbf{x}', t)$  durch die Entwicklung der Gleichung (7.2) mittels Taylorreihe um  $I_1(\mathbf{x}', t_1)$  angenähert werden:

$$\xi(\mathbf{x}', t) = I_1(\mathbf{x}', t_1) - I_2(\mathbf{x}', t_1) + \frac{\partial I_1}{\partial x'} dx' + \frac{\partial I_1}{\partial y'} dy' + \frac{\partial I_1}{\partial t} dt + \dots \quad (7.3)$$

Die Glieder quadratischer und höherer Ordnung werden bei der Reihenentwicklung vernachlässigt. Aus den Gleichungen (7.2) und (7.3) folgt

$$\frac{\partial I_1}{\partial x'} dx' + \frac{\partial I_1}{\partial y'} dy' + \frac{\partial I_1}{\partial t} dt = 0 \quad (7.4)$$

bzw.

$$\frac{\partial I_1}{\partial x'} \frac{dx'}{dt} + \frac{\partial I_1}{\partial y'} \frac{dy'}{dt} + \frac{\partial I_1}{\partial t} \underbrace{\frac{dt}{dt}}_{=1} = 0 \quad (7.5)$$

Die räumlichen und zeitlichen Ableitungen der Intensität  $\left(\frac{\partial I}{\partial x'}, \frac{\partial I}{\partial y'}, \frac{\partial I}{\partial t}\right)$  sind direkt über z. B. einfache Gradientenoperatoren messbare Größen und werden wie folgt bezeichnet:

$$I_{x'} = \frac{\partial I}{\partial x'}, \quad I_{y'} = \frac{\partial I}{\partial y'}, \quad I_t = \frac{\partial I}{\partial t}, \quad (7.6)$$

Die zur Schätzung des Optischen Flusses notwendigen Komponenten des Geschwindigkeitsvektorfelds  $f(\mathbf{x}', t)$  treten in Gleichung (7.6) in folgender Form auf:

$$f_{x'} = \frac{dx'}{dt}, \quad f_{y'} = \frac{dy'}{dt} \quad (7.7)$$

Schließlich ergibt sich die Kontinuitätsgleichung des Optischen Flusses (*Optical Flow Constraint Equation*), auch als Helligkeitskonstanzgleichung (*Brightness Constancy Constraint Equation*, **BCCE**) bezeichnet, zu

$$\frac{dI}{dt} = I_{x'} f_{x'} + I_{y'} f_{y'} + I_t = 0 \quad (7.8)$$

bzw.

$$(\nabla I)^T \mathbf{f} + I_t = 0 \quad (7.9)$$

mit

$$\nabla I = \begin{bmatrix} \frac{\partial I}{\partial x'} & \frac{\partial I}{\partial y'} \end{bmatrix} \quad \text{Räumlicher Bildgradient im Bildpunkt}$$

[Haußecker und Spies \(1999\)](#) sowie [Jähne \(2002\)](#) diskutieren die Modellierung von Bewegungsvektorfeldern ausführlich.

## 7.2 Bildzuordnung

Das Ziel der Bildzuordnung (*Matching*) ist die Detektion homologer Punkte bzw. Merkmale in verschiedenen Bildern und die anschließende Speicherung der zugehörigen Koordinaten. Die Herstellung von Zwei- oder Mehrbildkorrespondenzen ist ein Schwerpunkt in der digitalen Photogrammetrie und stellt ein zentrales Problem im Bereich der *Computer Vision* dar.

Die Bildzuordnung spielt traditionell bei der Orientierung von (Luft-)Bildpaaren insbesondere eine Rolle ([Baltsavias 1999](#)). Hierbei werden meist statische Szenen von unterschiedlichen Standpunkten sowie aus verschiedenen Blickrichtungen beobachtet und anschließend markante Punkte in den Bildern gemessen (räumliche Bildzuordnung). Aus photogrammetrischer Sicht ist eine Ausdehnung auf bewegungsanalytische Anwendungen (zeitliche Bildzuordnung) und die damit verbundene Adaption von bis dato nur zur Auswertung geo-spezifischer Sachverhalte genutzter Bildzuordnungsverfahren naheliegend. Dies wurde auch durch das Aufkommen digitaler Mess- und Auswertetechnik seit Anfang der neunziger Jahre des letzten Jahrhunderts bestärkt. Die Verbindung von strengem geometrisch-physikalischen Modellieren, präzisiertem Messen und automatischem Auswerten ermöglichte der Photogrammetrie die Ausweitung ihrer Kernkompetenzen auf die Registrierung und Auswertung kinematischer Vorgänge auf der Basis von bin-, trin- oder multiokularen Zuordnungsverfahren, welche seither in allen ingenieurtechnischen Bereichen und darüber hinaus in bspw. folgenden Gebieten zur Anwendung kommen ([McGlone u. a. 2004](#)):

- Bewegungsschätzungen aus Bildern
- Autonome Fahrzeug- und Flugzeugnavigation
- Robotersteuerung
- Qualitätskontrolle
- Veränderungsdetektion im Umweltbereich
- Menschliche Bewegungs- und Verhaltensanalyse

Die Zuordnung homologer Bildpunkte kann nicht auf der Basis des Grauwerts eines einzelnen Bildelements erfolgen, da dieser zumeist mehrfach vorhanden und eine eindeutige Zuordnung in einem zweiten Bild nicht möglich ist. Diese Problematik findet sich auch in der im vorherigen [Abschnitt 7.1.4](#) mathematisch beschriebenen Modellierung des zu schätzenden Bewegungsvektorfelds

wieder. Die BCCE (7.9) enthält  $n$  Vektorkomponenten im  $nD$  Raum. Bereits bei  $n = 2$  Dimensionen kann der Optische Fluss  $f(\mathbf{x}', t)$  nicht eindeutig geschätzt werden, da für jedes Bildelement nur eine Gleichung vorliegt. Lediglich der parallel zum Helligkeitsgradienten verlaufende Anteil des Bewegungsvektors ist eindeutig bestimmbar (Blendenproblem; vgl. Abschnitt 7.3.1). Es handelt sich also um ein klassisch unterbestimmtes Problem, welches zur Lösung der Aufgabe zusätzliche Beziehungen benötigt. Solche Bedingungen sind zum Beispiel eine gleichförmige Beleuchtung, eine Lambertsche Objektoberfläche oder eine reine Translation parallel zur Bildebene. In der Praxis sind diese Bedingungen niemals vollständig gegeben. Anstelle dessen wird unterstellt, dass die Bedingungen lokal in der Szene und somit auch lokal in der Bildebene erfüllt werden. Praktisch umgesetzt wird dies z. B. über die Ausnutzung von Informationen einer lokalen im betrachteten Bildpunkt  $\mathbf{x}'$  zentrierten Nachbarschaft  $\Sigma$ .

Für die Berechnung des Optischen Flusses existieren zahlreiche Methoden. Diese werden in der digitalen Bildverarbeitung häufig in differentielle Verfahren erster und zweiter Ordnung, Tensormethoden, blockweise arbeitende Korrelationsverfahren und Phasenkorrelationsverfahren unterteilt (Jähne 2002; Jähne u. a. 2007):

Differentielle Methoden gelten in der *Computer Vision* als klassischer Ansatz der Bewegungsanalyse zwischen konsekutiven Bildern. Sie bestimmen unter der Annahme einer kontinuierlich (differenzierbaren) Bildebene den Grauwertfluss, also die Änderung des Intensitätsverlaufs innerhalb einer Bildsequenz. Lokale und globale Methoden des Optischen Flusses basieren auf der Kontinuität des Grauwertflusses in Raum und Zeit (Lucas & Kanade 1981; Horn & Schunck 1981; Abschnitt 7.5). Sie betrachten somit die gesamte Bildsequenz und nicht nur diskrete Bildpaare. Als zusätzliche Bedingung wird eine globale Glättungsbedingung eingeführt. Für die lokal aufgestellte Methode wird im einfachsten Fall weiterhin angenommen, dass der Optische Fluss innerhalb eines kleinen Gebiets gleich groß ist, also lediglich konstante Translationen auftreten. Differentielle Methoden erlauben prinzipiell keine Veränderung der Lichtintensität. Neben der lokalen und globalen Methode zur Schätzung des Optischen Flusses zählt bspw. der Kanade-Lucas-Tomasi *Feature Tracker* zu den differentiell arbeitenden Bildzuordnungsverfahren (KLT-Tracker; Tomasi & Kanade 1991; Shi & Tomasi 1994).

Korrelationsbasierte Ansätze werden traditionell zur Verarbeitung von Stereobildern, also zur Analyse von Verschiebungen zwischen dem linken und dem rechten Bild einer Zweibildkonfiguration verwendet (Jähne 2002). Im Gegensatz zu den auf der Kontinuität des Optischen Flusses basierenden differentiellen Verfahren sind Korrelationsansätze relativ unempfindlich gegenüber radiometrischen Unterschieden und daher besonders für räumliche Bildzuordnungsaufgaben geeignet. Korrelationsbasierte Ansätze arbeiten blockweise und bestimmen innerhalb eines bestimmten Bildbereichs die Position der optimalen Ähnlichkeit eines Muster- und eines Suchfensters (*Template* und *Slave*). Korrelationsmethoden erreichen hohe (Subpixel-)Genauigkeit (Luhmann 2003), sind aber vergleichsweise rechenintensiv und im Allgemeinen auf die zeitlich diskrete Auswertung von zwei aufeinanderfolgenden Bil-



den beschränkt. Am bekanntesten ist die Kreuzkorrelation, welche den normierten Kreuzkorrelationskoeffizienten als Ähnlichkeitsmaß zwischen zwei Suchbereichen maximiert (Abschnitt 7.4). Alternativ minimiert die Kleinste-Quadrate-Anpassung (*Least Squares Matching, LSM*) die Summe der Quadrate der Grauwertunterschiede zwischen zwei Bildausschnitten über die Bestimmung eines neben der Translation zusätzlich um Rotations- und Maßstabs- effekte sowie Helligkeits- und Kontrasteinflüsse erweiterten Transformationsparametersatzes (Abschnitt 7.6).

Tensor- und Phasenkorrelationsmethoden sind für diese Arbeit nicht von Bedeutung.

Ebenso kann eine dem Photogrammeter vertraute Unterscheidung in merkmals- und flächenbasierte sowie relationale Verfahren erfolgen (Luhmann 2003):

Merkmalsbasierte Verfahren unterteilen sich in zwei Stufen: Zunächst werden durch Techniken der Merkmalsextraktion markante und möglichst einzigartige Bildstellen, z. B. Punktmuster, Kanten oder Ecken, über z. B. *Interest*-Operatoren in allen Bildern identifiziert (z. B. Förstner & Gülch 1987). Im weiteren Verlauf wird anhand ihrer Kenngrößen die Korrespondenz zwischen diesen punkt-, linien- oder selten auch flächenhaften Grauwertmustern hergestellt (*Interest*-Wert). Als Ergebnis liegt nicht nur ein einzelner Verschiebungsvektor, sondern ein Verschiebungsvektorfeld des gesamten Merkmals vor. Die Vorteile der merkmalsgestützten Zuordnung liegen in einem besseren Umgang mit Oberflächendiskontinuitäten sowie einer vereinfachten Bildanalyse durch die vorherige Extraktion interpretierbarer Merkmale. Im Vergleich zu flächenbasierten Methoden ist dieser Ansatz allerdings rauschempfindlicher und weniger genau. Ebenso bereiten Bilder mit zu stark ausgeprägter oder sich wiederholender Textur Probleme bei einer eindeutigen Merkmalszuordnung (Abschnitt 7.3). Die merkmalsbasierte Zuordnung findet häufig, meist in Verbindung mit Kernlinienschnitten und Pyramidenverfahren, bei der Erstellung von digitalen Geländemodellen Anwendung (Krzystek 1991; Maas 1996b).

Flächenbasierte Verfahren ordnen korrespondierende Bildregionen einander optimal bezüglich einer gegebenen Kostenfunktion einander zu. Die Eingabewerte für ein solches Verfahren sind die Grauwerte selbst oder Funktionen der Grauwerte (z. B. Gradienten). Neben der Intensität können aber auch weitere Attribute, wie z. B. die Distanzen eines Entfernungsbilds einer 3D-Kamera, integriert werden (Kapitel 14). Die Bildregion weist eine beliebige Form auf, wobei sich quadratische Masken aufgrund der Reduktion von Rechenschritten als praktikabel erweisen. Für die Ausmaskierung unerwünschter Bildbereiche sind aber auch z. B. rechteckige oder kreisförmige Fenster möglich. Einfache Korrelationstechniken wie die Kreuzkorrelation maximieren ein Ähnlichkeitsmaß, die Kleinste-Quadrate-Korrelation minimiert ein Differenzmaß zwischen einem festen Fenster im ersten Bild (Musterfenster) und einem beweglichen Fenster im zweiten Bild (Suchfenster). Ist die zu suchende Struktur bekannt, kann das *Template* auch synthetisch generiert werden. Die flächenbasierten Verfahren zeichnen sich besonders durch ihre Zuverlässigkeit und hohe Genauigkeit im Subpixelbereich aus. Problematisch hingegen ist die korrekte Zuordnung bei starken geometrischen Unterschieden zwischen den Bildern,

schwacher Textur oder Oberflächendiskontinuitäten ([Abschnitt 7.3](#)).

Relationale Verfahren basieren auf der Ähnlichkeit von topologischen Relationen ([Shapiro & Halick 1987](#); [Vosselman 1992](#)). Dieser Ansatz ist in der Praxis wenig bedeutsam, da die Komplexität aufgrund der Verwendung von Baumsuchverfahren sehr hoch ist und die Genauigkeit nicht an die ersten beiden Verfahren anknüpft.

## 7.3 Probleme bei der Bildzuordnung

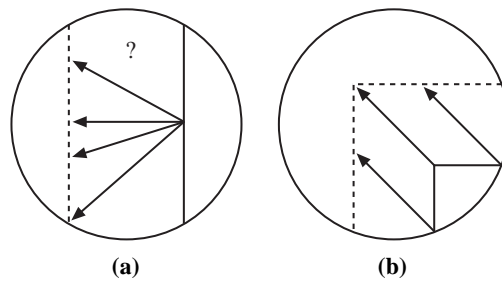
### 7.3.1 Blendenproblem

Die Bewegungsanalyse ist eng mit räumlichen und zeitlichen Grauwertänderungen verknüpft ([Abschnitt 7.1.1](#)). Beide Größen können über lokale Operatoren bestimmt werden (siehe u. a. [Abschnitt 7.2](#)). Solch ein Operator betrachtet bei der Analyse jedoch nur einen Ausschnitt des Objekts, als ob eine Blende über jenes gelegt wurde. Die [Abbildung 7.5a](#) zeigt als Beispiel hierfür eine Kante, die von der durchgezogenen zu der gestrichelten Linie parallel verschoben wurde. In diesem Fall ist die Translation, beschrieben durch einen Verschiebungsvektor, nicht eindeutig, da der Vektor von einem Punkt der Kante im ersten Bild zu einem beliebigen Punkt auf der verschobenen Kante im nächsten Bild zeigen kann. Es ist lediglich möglich, die senkrecht zur Kante liegende Komponente des Verschiebungsvektors zu bestimmen. Die parallele Komponente kann hingegen nicht eindeutig ermittelt werden. Diese Mehrdeutigkeit heißt Blendenproblem (*Aperturproblem*, *Aperture Problem*) und tritt bei allen lokal arbeitenden Bildzuordnungsverfahren auf. Eine eindeutige Bestimmung wird nur dann möglich, wenn die Operatormaske wie in [Abbildung 7.5b](#) so verschoben wird, dass sie die Ecke des Objekts einschließt.

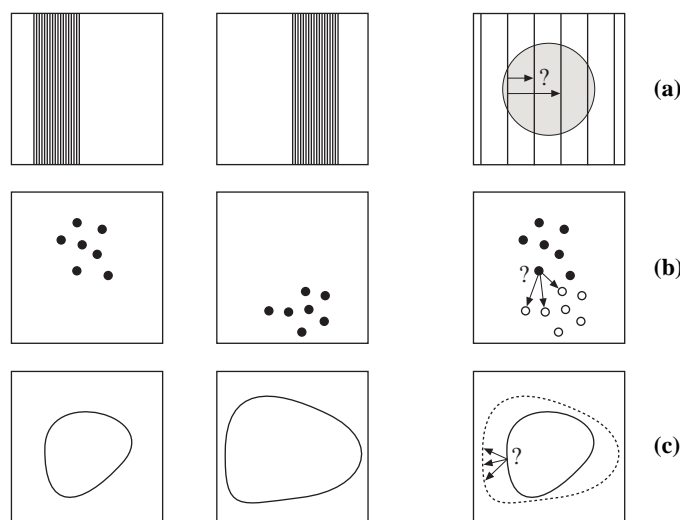
### 7.3.2 Allgemeines Korrespondenzproblem

Das oben beschriebene Blendenproblem ist ein Sonderfall des allgemeinen Korrespondenzproblems. Es beschreibt die Tatsache, dass ein Punkt in einem zweiten Bild aufgrund fehlender Unterscheidungsmöglichkeiten nicht eindeutig wiederzufinden ist. Die Lösung des Korrespondenzproblems ist eine der schwierigsten Aufgaben und noch immer ein allgemeines Ziel der Mustererkennung. Das grundlegende Problem der Bewegungsanalyse soll zunächst anhand einfacher Beispiele verdeutlicht werden.

Die [Abbildung 7.6a](#) zeigt die Translation eines hochfrequenten Musters (Gitternetz). Solange nur ein Ausschnitt des Objekts sichtbar ist, wäre auch eine Verschiebung um ein Vielfaches der Gitterweite denkbar und würde zu einer scheinbar richtigen Lösung führen. Das Problem hierbei ist, dass die reale (physikalische) Korrespondenz nicht mit der sichtbaren Korrespondenz im Bildausschnitt übereinstimmt. Obwohl also tatsächlich eine Korrespondenz besteht, kann diese visuell, aufgrund der periodischen Objekttextur, nicht bestimmt werden. Ein lokaler Zuordnungsalgorithmus wird mit hoher Wahrscheinlichkeit an einer falschen Stelle konvergieren. Das Korrespondenzproblem in dem



**Abb. 7.5:** Das Blendenproblem in der Bewegungsanalyse (nach Jähne 2002): (a) Mehrdeutigkeit der Verschiebungsvektoren an einer Kante. (b) Eindeutigkeit des Verschiebungsvektors an einer Ecke.



**Abb. 7.6:** Das Korrespondenzproblem bei (a) periodischen Strukturen, (b) nicht unterscheidbaren Teilchen und (c) deformierbaren Objekten (nach Jähne 2002).

eben diskutierten Beispiel kann analog auf nicht unterscheidbare Elemente oder deformierbare Körper angewendet werden (Abbildungen 7.6b und 7.6c).

Eine automatische Zuordnung korrespondierender Bildmerkmale nach einem der in Abschnitt 7.2 vorgestellten Verfahren ist, trotz intensiver Forschung seit den fünfziger Jahren, bis dato noch nicht vollständig erreicht. Auch existiert noch kein vollautomatisiertes, hoch präzises, zuverlässiges und effektives Verfahren, welches sich an verschiedene Bildinhalte anpassen kann. Die Wahrscheinlichkeit für Fehlkorrespondenzen wird in der Praxis durch die Aufstellung von Restriktionen, die die zu ermittelnden Verschiebungsvektoren erfüllen müssen, verringert. Beispiele hierfür sind die Richtung und die Größenordnung der erwarteten Bewegungen. Durch den Einsatz von Mehrbildtechniken können Zuordnungsaufgaben wesentlich erleichtert bzw. in einzelnen Fällen erst ermöglicht werden. Die Wahrscheinlichkeit von Mehrdeutigkeiten kann bspw. bei einem Stereoansatz durch die Verwendung eines dritten Bilds um mindestens eine Größenordnung reduziert werden (Maas 1997).

Im Folgenden werden weitere Beispiele als Sonderfälle des allgemeinen Korrespondenzproblems

betrachtet, welche sich vor allem im geodätisch-photogrammetrischen Kontext bewegen.

#### Geometrische und radiometrische Unterschiede zwischen den Bildern

Geometrische Unterschiede zwischen zwei konsekutiven Bildern entstehen bspw. durch Höhendifferenzen auf der Objektoberfläche, Änderungen der abgebildeten Objektgröße, Objektdeformationen oder auch durch Stellungs- und Standpunktabweichungen des Sensors. Im Zuge einer Bildzuordnung nach z. B. der **Methode der kleinsten Quadrate (MKQ)** können bereits mit dem gängigen Parametersatz einer 2D-Affintransformation geringe lineare geometrische Unterschiede modelliert werden (**Abschnitte 4.1** und **7.6**). Das Transformationsmodell kann ggf. um die Parameter einer Projektiv- oder Polynomtransformation erweitert werden (**Bethmann & Luhmann 2010**). Des Weiteren kommt es z. B. durch die Bewegung des Objekts zu Abweichungen in der Stärke der reflektierten und durch den Sensor aufgezeichneten Strahlung. Solche radiometrischen Unterschiede bewegen sich meist in geringen Größenordnungen und werden teilweise bereits im Zuordnungsmodell selbst berücksichtigt (normierter Kreuzkorrelationskoeffizient; radiometrische Anpassung während der Bildzuordnung nach der Methode der kleinsten Quadrate). Ebenso führen Variationen der Beleuchtungsverhältnisse zwischen den Aufnahmezeitpunkten zu Problemen bei der Bildpunktzuordnung, z. B. aufgrund des Sonnengangs stark variierenden Schattenwurfs. Solche Effekte können, ein Oberflächenmodell des zu verfolgenden Objekts vorausgesetzt, streng modelliert werden. Eine Reduktion des Einflusses geometrischer und radiometrischer Unterschiede ist auch durch die gezielte Wahl der zu verfolgenden Muster, durch die Einführung zusätzlicher Bedingungen oder die Vergrößerung der Muster-/Suchfenster möglich. Am Beispiel der mehrtägigen Analyse des Fließverhaltens eines Gletschers implementiert **Westfeld (2005)** z. B. zusätzliche räumliche und zeitliche Zwangsbedingungen, welche die zu verfolgenden Objektpunkte bzw. ihre Trajektorien erfüllen müssen. Das Korrespondenzproblem, verursacht durch z. B. den Übergang von Tag zu Nacht bzw. Schattenwürfe aufgrund der zerklüfteten Gletscheroberfläche, konnte so deutlich verringert werden.

#### Ungünstiges Signal-Rausch-Verhältnis

Da die Bildelemente lediglich anhand ihrer Grauwerte bzw. ihrer Nachbarschaftsbeziehungen lokalisiert werden, ist ein guter Bildkontrast für eine statistisch sichere Zuordnung unabdingbar. Ein niedriges **Signal-Rausch-Verhältnis (SNR)** verringert die Ähnlichkeit zwischen zwei korrespondierenden Bildausschnitten und somit die Erfolgsrate der Bildzuordnung.

#### Näherungswerte

Optional finden Näherungswerte zur Eingrenzung des Suchbereichs Verwendung. Beim Verfahren der Kreuzkorrelation kann, insofern das Bewegungsverhalten des zu verfolgenden Objekts annähernd bekannt ist, der Suchbereich durch die Definition geeigneter Prädiktionen eingegrenzt und somit die Wahrscheinlichkeit multipler Lösungen reduziert werden (**Abschnitt 7.4**). **Maas u. a. (2005)**

implementieren beispielsweise eine 2D-Kreuzkorrelation mit einem vorwärtsorientierten beschränkten Suchbereich zur Bestimmung raumzeitlich aufgelöster Bewegungsfelder eines grönländischen Gletschers. Iterative Verfahren wie LSM (Abschnitt 7.6) benötigen aufgrund der Linearisierung der aufgestellten Modellgleichungen zwingend Näherungswerte. Sind deren Prädiktionen zu schlecht, konvergiert der Algorithmus in der Regel nicht. Eine Erhöhung des Konvergenzradius kann hier z. B. durch die hierarchische Anwendung des Zuordnungsverfahrens auf verschiedene Auflösungsstufen einer Bildpyramide und durch die damit einhergehende sukzessive Verbesserung der Näherungen erreicht werden. In Verbindung mit flächenbasierten Verfahren werden Bildpyramidentechniken in der Photogrammetrie meist bei der automatischen relativen Orientierung von Luftbildern (z. B. Liang & Heipke 1996) oder bei der Generierung von digitalen Geländemodellen (z. B. Baltsavias & Stallmann 1992) eingesetzt. In der Bewegungsanalyse kommen sie z. B. bei der Bestimmung von Wasserströmungen aus Satellitenbildpaaren (Hasanlou & Saradjian 2006) oder bei der markerlosen Verfolgung von Bekleidungsstücken (*Garment Tracking*; Hilsmann & Eisert 2008) zum Einsatz.

#### Fehlende Objektteile

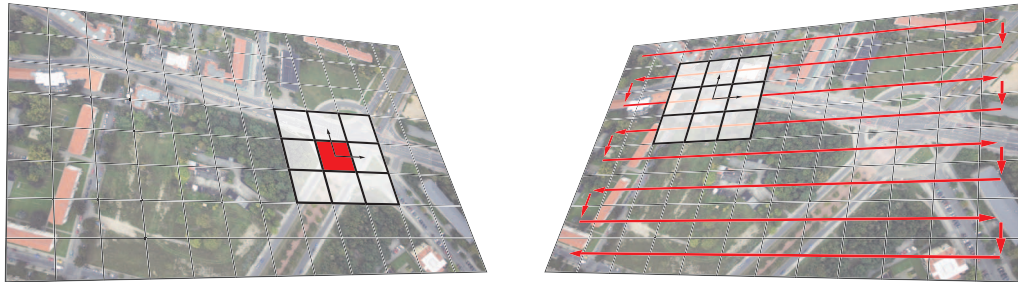
Partielle und/oder temporäre Verdeckungen von in Bildsequenzen zu verfolgenden Objekten werden in der Regel durch das begrenzte Sichtfeld der Kameras sowie durch andere Objekte im Vordergrund hervorgerufen. Im Sinne einer vernünftigen automatischen Auswertung von langen Bildsequenzen, sollten, insbesondere bei einer flächenbasierten Zuordnung von Bild zu Bild, auftretende Verdeckungen zumindest detektiert werden, um ein interaktives Eingreifen eines Operators aufgrund des Drifts des Musterfensters zu minimieren. Maas (1992a) und Putze (2004) lösen bei der Bestimmung dreidimensionaler Partikeltrajektorien in Gas- oder Flüssigkeitsströmungen durch Verdeckungen induzierte Korrespondenzprobleme durch die Verwendung zusätzlicher Kameras. Hier können mittels redundanter Mehrbildaufnahmen fehlende Zuordnungen durch Kombinationen gültiger Zuordnungen ersetzt werden. Einen anderen Ansatz wählt Schulze (2008): Am Beispiel der Blickerfassung (*Eye Tracking*) wurden automatisch die Zeitpunkte detektiert, an denen ein Auge durch seinen Lidschluss verdeckt blieb. Die Position des verdeckten Auges wurde näherungsweise über sichtbare Merkmale wie die Augenbrauen oder die Nase geschätzt. Diese Schätzung war notwendig, um den durch eine Verdeckung hervorgerufenen Drift des Suchfensters zu korrigieren. Die fehlenden Positionen während des Lidschlusses konnten linear interpoliert werden. Als Alternative ist eine Prädiktion der verdeckten Bewegung auf Grundlage eines Kalman-Filter-Ansatzes möglich (Bozek 2010).

## 7.4 Kreuzkorrelation

Die Kreuzkorrelation als flächenbasiertes Korrelationsverfahren bestimmt die Position der größten Ähnlichkeit eines Referenzmusters in einem Suchbild. Das Maß für die Ähnlichkeit zweier Muster (Korrelationsmaß) ist der Kreuzkorrelationskoeffizient, welcher zu maximieren ist.

Ein Suchfenster wird sukzessiv mäanderförmig über das Suchbild bewegt (Abbildung 7.7), wobei

die Größe des einbezogenen Bildbereichs in Abhängigkeit der Größenordnung der erwarteten Bewegung und der Bewegungsrichtung eingeschränkt werden kann. Der Kreuzkorrelationskoeffizient als zu maximierendes Ähnlichkeitsmaß wird an jeder Verschiebungsposition aus den Bildinformationen des Suchfensters berechnet und mit den entsprechenden Bildkoordinaten des zentralen Bildelements gespeichert.



**Abb. 7.7:** Prinzip der Kreuzkorrelation. Links: Ein Musterbild  $I_1$  mit einer um den zuzuordnenden Bildpunkt  $\mathbf{x}'$  zentrierten Nachbarschaft  $g_1$ . Rechts: Das Suchfenster  $g_2$  wird mäandrierend über das Bild  $I_2$  verschoben, um den räumlich oder zeitlich korrespondierenden Bildpunkt  $\mathbf{x}' + \Delta\mathbf{x}'$  zu lokalisieren.

Das Musterfenster der Größe  $(R \times C)$  px sei bezeichnet als  $g_1$  mit seinem Bildinhalt  $g_1(x', y')$ , das Suchfenster analog als  $g_2$ . Da sowohl bei einer räumlichen als auch bei einer zeitlichen Bildzuordnung radiometrische Unterschiede zwischen den beiden Bildfenstern auftreten können, wird der normierte Kreuzkorrelationskoeffizient direkt eingeführt:

$$\rho(\Delta\mathbf{x}') = \frac{\sum_{x'=0}^{C-1} \sum_{y'=0}^{R-1} (g_1(\mathbf{x}') - \bar{g}_1) \cdot (g_2(\mathbf{x}' + \Delta\mathbf{x}') - \bar{g}_2)}{\sqrt{\sum_{x'=0}^{C-1} \sum_{y'=0}^{R-1} (g_1(\mathbf{x}') - \bar{g}_1)^2} \cdot \sqrt{\sum_{x'=0}^{C-1} \sum_{y'=0}^{R-1} (g_2(\mathbf{x}' + \Delta\mathbf{x}') - \bar{g}_2)^2}} \quad (7.10)$$

mit

$\mathbf{x}'(x', y')$	Koordinaten eines lokalen Koordinatensystems, gelagert in der oberen linken Fensterecke
$\Delta\mathbf{x}'(\Delta x', \Delta y')$	Relative Translationen in $x'$ und $y'$
$\bar{g}_1, \bar{g}_2$	Arithmetische Grauwertmittel

Das Ähnlichkeitsmaß  $\rho$  ist unempfindlich gegenüber Beleuchtungsänderungen. Mögliche Helligkeitsunterschiede werden in Gleichung (7.10) über eine Grauwertzentrierung berücksichtigt, was aus mathematischer Sicht der Kovarianz zwischen den Grauwerten der zwei Fenster entspricht. Das Produkt der Standardabweichungen der Grauwerte des jeweiligen Fensters ist der Nenner, welcher das Verfahren unempfindlich gegenüber Kontrastunterschieden macht.

Der Wertebereich des normierten Kreuzkorrelationskoeffizienten erstreckt sich von  $-1$  bis  $1$ . Ein Korrelationswert von Null lässt auf keinerlei Beziehungen zwischen den Fensterinhalten schließen, wohingegen identische Fenster den Wert  $1$  und Negativkorrelationen den Wert  $-1$  aufweisen. In der Praxis bewegen sich die Kreuzkorrelationskoeffizienten zwischen  $0,6$  und  $1,0$ . Zahlenwerte unter-



halb dieser Größenordnungen lassen auf eine falsche Bildzuordnung schließen. Bei starkem Bildrauschen oder Verdeckungen kann jedoch auch eine korrekte Zuordnung geringe Ähnlichkeitsmaße hervorrufen. Im umgekehrten Fall bewirken zum Beispiel multiple Lösungen ebenfalls ausgeprägte Kreuzkorrelationskoeffizienten. In der Nachbarschaft des eigentlich gesuchten Korrelationspeaks entstehen weitere Nebenmaxima mit unter Umständen gleichen Korrelationskoeffizienten (vgl. z. B. Westfeld 2005).

Das Ergebnis einer Bildzuordnung nach der Methode der Kreuzkorrelation ist eine Matrix von Kreuzkorrelationskoeffizienten mit maximal der Größe der Eingangsbilder. Werden diese Werte als Funktion von  $x'$  und  $y'$  aufgetragen, ist die Stelle der maximalen Korrelation mit größter Wahrscheinlichkeit die Position des Referenzmusters im Suchbild. Im Anschluss kann um die Stelle des Maximums ein Paraboloid über z. B. eine Ausgleichung nach der MKQ approximiert werden (Abschnitt 4.4.1), um die gesuchte Position im Subpixelbereich zu ermitteln. Dies ist notwendig, da die oben ermittelte Maximaposition eine Integerkoordinate ist und somit nicht zwangsläufig die tatsächliche Stelle der maximalen Korrelation widerspiegeln muss.

Die Kreuzkorrelation ist ein robustes Verfahren, welches sich besonders durch die in der Regel einfache Implementierung in ein Computerprogramm auszeichnet. Die erreichbare Bildmessgenauigkeit liegt bei 0,1 px (Baltasvias 1999). Die wichtigsten Probleme der Kreuzkorrelation sind der hohe Rechenaufwand, die Nichtberücksichtigung geometrischer Verzerrungen sowie die fehlerfreie Zuordnung bei vorhandenen Nebenmaxima.

Weitere Erläuterungen finden sich bspw. in (Ballard & Brown 1982; Nevatia 1982; Luhmann 2003; McGlone u. a. 2004).

## 7.5 Methoden des Optischen Flusses

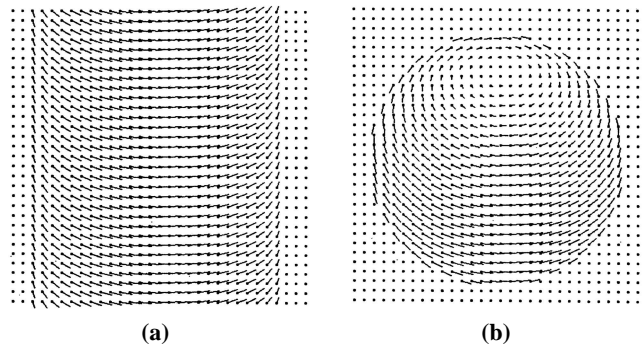
Auf den Gebieten der *Computer Vision* gibt es eine Vielzahl von Methoden zur Berechnung des Optischen Flusses (Abbildung 7.8), wobei die meisten auf der differentiellen Methode aufbauen (Barron u. a. 1994). Die differentielle Form der Kontinuitätsgleichung soll gelöst und so die Bildgeschwindigkeit unter der Annahme einer kontinuierlichen Bildoberfläche über räumliche und zeitliche Ableitungen der Intensität berechnet werden (vgl. Gleichungen 7.8 bzw. 7.9):

$$\nabla I(x', y', t) \cdot (f_{x'}, f_{y'}, 0) + I_t(x', y', t) = 0 \quad (7.11)$$

Im Folgenden werden zwei Verfahren, ein lokaler und ein globaler Ansatz, zur Schätzung des Optischen Flusses kurz vorgestellt. Beide basieren auf der Gleichung (7.11). Für eine Vertiefung der Thematik sei auf z. B. (Haußecker & Spies 1999; Jähne 2002) verwiesen.

Die Implementierung von Methoden des Optischen Flusses sind ebenfalls zahlreich. Black u. a. (1997) verfolgen und rekonstruieren menschliche Bewegungen über ein hierfür speziell parametrisiertes Modell des Optischen Flusses. Eine auf dem Optischen Fluss basierende Überwachung des toten Winkels in Automobilanwendungen wird in (Batavia u. a. 1997) vorgestellt. Garbe u. a. (2008)

verwenden sie bspw., um Mikroströmungen in der Strömungsmechanik zu messen.



**Abb. 7.8:** Optisches Flussfeld eines Zylinders (a) und einer Kugel (b) (Horn & Schunck 1981).

### 7.5.1 Lokale Methoden

Das in der Praxis am häufigsten verwendete lokale Modell der Geschwindigkeitsschätzung wurde von Lucas und Kanade (1981) entwickelt. Unter der Annahme konstanter Geschwindigkeitsvektoren in einer kleinen Nachbarschaft  $g$  um das zu untersuchende Bildelement  $\mathbf{x}'$  kann die BCCE (7.9) des Optischen Flusses für jedes, sich innerhalb dieser Umgebung befindende Bildelement aufgestellt werden. Diese Forderung führt zu einem überbestimmten Gleichungssystem und ermöglicht die Schätzung der zwei unbekannt Geschwindigkeitsvektoren  $\mathbf{f}(f_{x'}, f_{y'})$  unter der Minimierung von

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \int_{-\infty}^{\infty} w(\mathbf{x}') [(\nabla g)^T \mathbf{f} + g_t]^2 d\mathbf{x}' \quad (7.12)$$

Die Gewichtsfunktion  $w(\mathbf{x}')$  besteht in der Regel aus den Koeffizienten eines 2D-Gauß-Filters, um den Einfluss des Centerpixels im Vergleich zur umgebenden Nachbarschaft zu erhöhen.

Im Rahmen einer gewichteten MKQ (Abschnitt 4.4.2) kann der Optische Fluss geschätzt werden:

$$\begin{aligned} \mathbf{f} &= \mathbf{A}^{-1} \mathbf{b} \\ &= \begin{bmatrix} g_{x'x'} & g_{x'y'} \\ g_{x'y'} & g_{y'y'} \end{bmatrix}^{-1} \begin{bmatrix} g_{tx'} \\ g_{ty'} \end{bmatrix} \end{aligned} \quad (7.13)$$

Das Gleichungssystem (7.13) kann nur gelöst werden, insofern die Inverse zu  $\mathbf{A}$  existiert, also z. B.  $\det(\mathbf{A}) \neq 0$  gilt. Dies ist nicht der Fall, wenn alle räumlichen Ableitungen in der Nachbarschaft gleich Null sind (uniformer Bildinhalt; fehlende Textur) oder alle Gradienten in dieselbe Richtung zeigen (Textur in nur einer Koordinatenrichtung). Im Abschnitt 7.3 wird dieses Korrespondenzproblem ausführlich behandelt.

Lokale Ansätze sind bekannt für ihr robustes Verhalten gegenüber Rauschen. Sie unterstützen



die Berechnung des Optischen Flusses für wenige Bildpunkte, wohingegen die im folgenden **Abschnitt 7.5.2** behandelten globalen Techniken einhundert Prozent dichte Flussfelder bestimmen und dementsprechend mehr Rechenleistung in Anspruch nehmen.

### 7.5.2 Globale Methoden

Globale Methoden wie die von **Horn und Schunck (1981)** fordern keine räumlich lokale Kohärenz innerhalb eines kleinen Bildausschnitts, sondern eine globale über den gesamten Bildbereich. Das heißt, dass die **BCCE (7.9)** des Optischen Flusses über  $I$  zu minimieren ist:

$$\hat{\mathbf{f}} = \arg \min_f \int_I [(\nabla I)^T \mathbf{f} + I_t]^2 d\mathbf{x}' + \lambda^2 \|\mathbf{e}_S^2\| \quad (7.14)$$

mit

$$\|\mathbf{e}_S^2\| = \int_I \|\nabla f_{x'}(\mathbf{x}')\|^2 + \|\nabla f_{y'}(\mathbf{x}')\|^2 d\mathbf{x}'$$

Zur Vermeidung eines schlecht konditionierten Gleichungssystems wird eine Nebenbedingung in Form eines Regularisierungsterms  $\|\mathbf{e}_S^2\|$  eingeführt, welche eine globale Glättung fordert und eine langsame Variation des Strömungsfelds unterstellt. Dieser zusätzliche *Smoothness Constraint* bewirkt, dass alle Bewegungen weich und gleichmäßig verlaufen und somit die Grauwerte über den gesamten Bildbereich differenzierbar sind. Der Parameter  $\lambda$  beeinflusst hierbei die Auswirkung des Glättungsterms (je größer  $\lambda$ , desto glatter das Flussfeld). An Bildstellen mit hohen räumlichen Grauwertgradienten ( $\nabla I \gg 0$ ) dominiert der erste Summand in Gleichung (7.14). Die Berechnung des Optischen Flusses folgt also der Kontinuitätsgleichung (7.9). An Stellen mit gleichgerichteten oder keinen Bildgradienten können über den gewichteten Glättungsterm Informationen aus der Nachbarschaft des zu betrachtenden Pixels erhalten werden (*Filling-in Effect*). Dies führt zu einem dichten Flussfeld und macht weitere zusätzliche Informationen überflüssig.

Die Minimierung von Gleichung (7.14) geschieht iterativ und ergibt eine Approximation des Optischen Flusses  $\mathbf{f}^{(k+1)}$  zum Zeitpunkt  $k + 1$ :

$$\mathbf{f}^{(k+1)} = \bar{\mathbf{f}}^{(k)} - \nabla I \frac{\nabla I \cdot \bar{\mathbf{f}}^{(k)} + I_t}{\|\nabla I\|^2 + \lambda^2} \quad (7.15)$$

mit

$$\bar{\mathbf{f}}^{(k)} \quad \text{Lokales Mittel von } \mathbf{f}^{(k)}$$

**Barron u. a. (1994)** beobachten, dass globale Methoden, im Vergleich zu lokalen Techniken, rauschanfälliger sind. Eine mögliche Erklärung für dieses Verhalten liegt darin, dass Rauschen große Bildgradienten, welche in Gleichung (7.14) als Gewichte dienen, hervorruft. Das konstante Gewicht  $\lambda$  des Glättungsterms verursacht an Positionen mit großen Bildgradienten weniger glättende Wirkung als an anderen Stellen. In Konsequenz daraus sind Flussfelder an verrauschten Bildstrukturen

weniger regularisiert. Diese Sensibilität ist der Nebeneffekt des begehrten *Filling-in Effects*. Eine Vergrößerung des Regularisierungsparameters  $\lambda$  würde das optische Strömungsfeld in verrauschten Regionen glätten, allerdings auch die Eliminierung kleiner Details in flacheren Bildbereichen herbeiführen. Ein wesentlicher Vorteil der globalen Algorithmen ist, dass das gesamte Flussfeld bestimmt wird.

## 7.6 Least Squares Matching

In der Photogrammetrie wird häufig das flächenbasierte Korrelationsverfahren nach der Methode der kleinsten Quadrate (*Least Squares Matching, LSM*) zur Extraktion von Bewegungsinformationen aus Bildsequenzen angewendet (Ackermann 1984; Grün 1985; Förstner 1986). *LSM* ist ein Verfahren der Bildzuordnung, das homologe Bildbereiche durch eine Affintransformation mit dem Ziel der Minimierung der Summe der Quadrate der Grauwertdifferenzen bestimmt. Bei guten Näherungswerten und bandbreiten-begrenztem Bildsignal konvergiert das Verfahren nach wenigen Iterationen. *LSM* weist eine Messgenauigkeit von etwa 0,01 px bis 0,1 px auf und liefert als stochastisches Verfahren zusätzlich zur eigentlichen Bewegungsinformation Parameter zur Beschreibung der Genauigkeit und Zuverlässigkeit der Bewegungsvektoren.

Aus traditionell photogrammetrischer Sicht wird *LSM* für räumliche Zuordnungsaufgaben bei bspw. der Herstellung von Stereoluftbildkorrespondenzen und der anschließenden Geländemodellerstellung verwendet (z. B. Helava 1988; vgl. auch Abschnitt 7.2). Für die Herstellung zeitlicher Korrespondenzen bei Bewegungsvorgängen ist in der Literatur auch die Begrifflichkeit *Least Squares Tracking (LST)* gebräuchlich (z. B. D'Apuzzo u. a. 2000).

Das Prinzip der Bildzuordnung nach der Methode der kleinsten Quadrate wird im Abschnitt 7.6.1 am Beispiel eines *2D-LSM* erläutert, welches ursprünglich als ein Verfahren zur subpixelgenauen Bestimmung zweidimensionaler Informationen aus Bilddaten entwickelt wurde. Darüber hinaus existieren zahlreiche Erweiterungen und Adaptionen, bspw. die spatiotemporale Zuordnung dreidimensionaler Datensätze (Abschnitt 7.6.2) oder die Übertragung des klassischen Ansatzes auf Oberflächenpunktvolken (Abschnitt 7.6.3). Nicht zuletzt wird im Rahmen dieser Arbeit ein integrierter *2,5D-LST*-Ansatz zur Auswertung von *3D*-Kamerabildsequenzen präsentiert (Kapitel 14).

### 7.6.1 2D Least Squares Matching

Bei dem Verfahren der Bildzuordnung nach der Methode der kleinsten Quadrate werden die Beziehungen zwischen den Grauwerten zweier Bilder  $I_1$  und  $I_2$  innerhalb eines Musterfensters  $g_1$  und eines Suchfensters  $g_2$  als nichtlineare Verbesserungsgleichungen in einer vermittelnden Ausgleichung formuliert (Abschnitt 4.4.2). In Analogie zu den in Abschnitt 7.1.4 gelegten Grundlagen zur Modellierung eines optischen Strömungsfelds, kann die folgende Beziehung für bis auf einen kleinen

Rauschanteil  $e(x', y')$  identische Bildausschnitte aufgestellt werden:

$$g_1(x', y') - e(x', y') = g_2(x', y') \quad (7.16)$$

Vor allem bei der Zuordnung statischer Szenen unterschiedlicher Zeitpunkte, aber auch bei kinematischen Vorgängen innerhalb einer nicht vollständig kontrollierbaren Experimentalumgebung, sind radiometrische Variationen zwischen den Bildausschnitten nicht auszuschließen und müssen modelliert werden:

$$g_1(x', y') - e(x', y') = r_0 + r_1 \cdot g_2(x', y') \quad (7.17)$$

Zur Bestimmung der radiometrischen Parameter  $r_0$  (additiver Helligkeitsterm) und  $r_1$  (multiplikativer Kontrastterm) existieren zwei verschiedene Philosophien: Sie können entweder vor der eigentlichen Ausgleichung über eine lineare Histogrammanpassung bestimmt oder im Zuge des Ausgleichungsansatzes simultan mitgeschätzt werden. Der Vorteil der ersten Variante liegt darin, dass eine theoretisch vollkommene Unabhängigkeit der Bewegungsschätzung von radiometrischen Einflüssen erzielt wird (Baltsavias 1991).

Bei genügend kleinen Bildausschnitten kann die geometrische Transformation durch eine ebene Affintransformation beschrieben werden (Abbildung 7.9; Abschnitt 4.1). Die sechs Parameter einer affinen Abbildung erlauben die Berücksichtigung von relativen Translationen in  $x'$ - und  $y'$ -Richtung ( $a_0, b_0$ ), beliebigen Maßstäben in beiden Koordinatenrichtungen ( $a_1, b_2$ ) sowie Dreh- und Scherungswinkeln zwischen den Koordinatenachsen ( $a_2, b_1$ ):

$$\begin{aligned} x'_2 &= x'_1 + \Delta x' = a_0 + a_1 x'_1 + a_2 y'_1 \\ y'_2 &= y'_1 + \Delta y' = b_0 + b_1 x'_1 + b_2 y'_1 \end{aligned} \quad (7.18)$$

Die oben aufgestellte ursprüngliche Beobachtungsgleichung (7.17) muss aufgrund der Nichtlinearität der Bildfunktion  $g_1$  durch eine Reihenentwicklung nach Taylor an der Stelle von Näherungswerten für alle unbekanntem Transformationsparameter linearisiert werden:

$$\begin{aligned} g_1(x', y') - e(x', y') &= g_2^0(x', y') + g_{x'} da_0 + g_{x'} x' da_1 + g_{x'} y' da_2 \\ &\quad + g_{y'} db_0 + g_{y'} x' db_1 + g_{y'} y' db_2 \\ &\quad + r_0^0 + r_1^0 g_2^0(x', y') \end{aligned} \quad (7.19)$$

mit

$$g_{x'} = \partial g_2^0(x', y') / \partial x', \quad g_{y'} = \partial g_2^0(x', y') / \partial y'$$

Bedingung für die Konvergenz des Algorithmus ist, neben der Notwendigkeit von Textur mit Gradienten in beiden Bildrichtungen, die Bereitstellung genügend genauer Näherungswerte (vgl. Abschnitt 7.3). Bei einer simultanen radiometrischen und geometrischen Anpassung werden die radiometrischen Parameter meist mit  $r_0^0 = 0$  und  $r_1^0 = 1$  initialisiert. Die affinen Rotations- und Scherungs-

parameter  $(a_2, b_1)^0$  werden bei zeitlichen Bildzuordnungen in der Praxis häufig mit Null, die Maßstabsparameter  $(a_1, b_2)^0$  mit Eins approximiert. Für konsekutive Bilder einer Sequenz wird hier von sehr geringen Fensterdeformationen innerhalb des Konvergenzradius des Algorithmus ausgegangen, also innerhalb der halben Fenstergröße von der bestmöglichen Lösung. Anders verhält es sich bei den Translationsparametern  $(a_0, b_0)^0$ : Treten schnelle Bewegungen auf und/oder wird die Szene mit einer geringen zeitlichen Auflösung videografiert, reicht die Initialverschiebung des Suchfensters auf die Stelle des Musterfensters nicht aus. Die Bereitstellung verbesserter Näherungswerte kann dann auf unterschiedliche Art und Weise erfolgen:

- Ein hierarchischer Pyramidenansatz erlaubt, in Abhängigkeit der Anzahl der Pyramidenstufen, eine sukzessive Vergrößerung des Konvergenzradius und folglich die Verwendung der Lösung des letzten Zeitschritts als Näherung (z. B. [Baltsavias 1988](#)).
- Der Konvergenzradius kann ebenfalls durch eine Vergrößerung der Fensterdimensionierung erreicht werden. Hierbei ist jedoch zu beachten, dass bei einer endlichen Vergrößerung das unterstellte geometrische Modell ggf. nicht mehr verwendet werden kann.
- Zuordnungsverfahren wie bspw. die Kreuzkorrelation ([Abschnitt 7.4](#)) oder die *Fast Fourier Transformation* (FFT) bestimmen Bildkorrespondenzen ohne einen Konvergenzradius im klassischen Sinne. Deren Lösung dient als sehr gute Näherung für ein nachgeschaltetes LSM (z. B. [Jähne 1999](#); [Westfeld 2005](#)).
- Die Implementierung eines Kalman-Filter-Ansatzes ermöglicht, insofern die zu erwartende Bewegung modelliert werden kann, eine Extrapolation des Trajektorienverlaufs (z. B. [Bozek 2010](#)).

Im Zuge einer Ausgleichung nach vermittelnden Beobachtungen werden die geometrischen und, insofern nicht bereits im Voraus bestimmt, auch die radiometrischen Parameter im GMM geschätzt:

$$\begin{aligned} \mathbf{v} &= \mathbf{A}\hat{\mathbf{x}} - \mathbf{l} \\ \hat{\mathbf{x}} &= (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{l} \end{aligned} \tag{7.20}$$

Gemäß der theoretischen Ausführungen des [Abschnitts 4.4.2](#) bezeichnet  $\mathbf{A}$  die Koeffizientenmatrix, gebildet aus den partiellen Ableitungen der Bildfunktion nach allen unbekanntem Parametern,  $\mathbf{l}$  den Beobachtungsvektor als Grauwertdifferenzvektor sowie  $\mathbf{v}$  und  $\hat{\mathbf{x}}$  den Verbesserungs- und Lösungsvektor. Da die Grauwertbeobachtungen in der Regel gleichartig und gleich genau sind, ist die Gewichtsmatrix  $\mathbf{P}$  eine Einheitsmatrix.

Das Gleichungssystem (7.20) ist zumeist hoch redundant: Den maximal acht unbekanntem Transformationsparametern stehen alle Grauwerte des Bildausschnitts als Beobachtungen entgegen. Es wird iterativ gelöst und konvergiert, unter der Voraussetzung hinreichend genauer Näherungswerte, nach wenigen Durchläufen. Als statistisches Schätzverfahren liefert LSM zusätzlich zum eigentlich interessierenden Lösungsvektor Fehlermaße, welche zur Beurteilung von Genauigkeit, Zuverlässigkeit und Signifikanz der Lösung herangezogen werden können ([Abschnitt 4.5](#)).

**2D-LSM** ist eines der genauesten Bildzuordnungsverfahren (Ackermann 1984; McGlone u. a. 2004) und kann in idealen Fällen für die aus photogrammetrischer Sicht relevanten Verschiebungsparameter Subpixelgenauigkeiten bis 0,01 px erreichen. Generell ist diese theoretische Genauigkeit zu optimistisch, da hier die stochastischen Messwerte zweier Bildausschnitte verglichen werden und nicht, wie in der Ausgleichsrechnung üblich, Messwerte mit einer analytischen Funktion (Abschnitt 4.4). Darüber hinaus führt die Berücksichtigung unmodellierter Fehler zu einer schlechteren empirischen Genauigkeit der Translationsbestimmung. Sie beläuft sich nach (Baltsavias 1999)

- für kontrastreiche und flache Zielmarken auf 0,01 – 0,1 px,
- für kontrastarme und flache Zielmarken auf 0,2 – 0,3 px,
- für natürliche Punkte auf allgemeinen Oberflächen auf 0,5 – 1,0 px.

### 7.6.2 3D Least Squares Matching

Das oben beschriebene **2D-LSM** kann über die Erweiterung um eine Dimension in einen dreidimensionalen Ansatz überführt werden (Maas u. a. 1994). Folglich arbeitet **3D-LSM** mit **3D-Volumendaten**, bestehend aus einzelnen Volumenelementen (*Volumetric Picture Element*; Voxel), anstelle von **2D-Bildern** und **Pixeln**. Die ebene Affintransformation mit sechs Parametern wird durch eine **3D-Affintransformation** mit 12 Parametern ersetzt (Abbildung 7.10):

$$\begin{aligned}x'_2 &= x'_1 + \Delta x' = a_0 + a_1 x'_1 + a_2 y'_1 + a_3 z'_1 \\y'_2 &= y'_1 + \Delta y' = b_0 + b_1 x'_1 + b_2 y'_1 + b_3 z'_1 \\z'_2 &= z'_1 + \Delta z' = c_0 + c_1 x'_1 + c_2 y'_1 + c_3 z'_1\end{aligned}\tag{7.21}$$

In Analogie zur Gleichung (7.16) ergibt sich für die raumzeitlich korrespondierende Kuboide  $g_1(x', y', z')$  bzw.  $g_2(x', y', z')$  einer Voxeldatensequenz der folgende Zusammenhang:

$$\begin{aligned}g_1(x', y', z') - e(x', y', z') &= g_2^0(x', y', z') \\ &+ \frac{\partial g_2^0(x', y', z')}{\partial x'} dx' + \frac{\partial g_2^0(x', y', z')}{\partial y'} dy' + \frac{\partial g_2^0(x', y', z')}{\partial z'} dz'\end{aligned}\tag{7.22}$$

mit

$$\begin{aligned}dx' &= (\partial x' / \partial dp_i) dp_i, \quad dy' = (\partial y' / \partial dp_i) dp_i, \quad dz' = (\partial z' / \partial dp_i) dp_i, \\ p_i &\in (a_0, a_1, a_2, a_3, b_0, b_1, b_2, b_3, c_0, c_1, c_2, c_3, r_0, r_1)\end{aligned}$$

Unter der Verwendung von Grauwertgradienten in drei Koordinatenrichtungen wird die Gleichung (7.22) an der Stelle der Näherungswerte linearisiert und für jedes Voxel als Beobachtungsgleichung in einer vermittelnden Ausgleichung aufgestellt. Die weitere Vorgehensweise gestaltet sich analog zu Abschnitt 7.6.1.

In (Maas 1993; Maas u. a. 1994) werden durch **3D-LSM** dreidimensionale Geschwindigkeitsfelder in chemischen Mischvorgängen bestimmt. Dracos u. a. (1995), Maas (1996a) und Kitzhofer u. a.

(2010) nutzen zusätzlich die Möglichkeit, die Transformationsparameter aus **3D-LSM** zur Analyse von auf den Tensor wirkenden Scherkräften, Deformationen und Rotationen bei Strömungsversuchen heranzuziehen. Schulze u. a. (2005) setzen **3D-LSM** zur Objektverfolgung in multitemporalen Computertomographiedatensätzen bei Sinterungsversuchen ein.

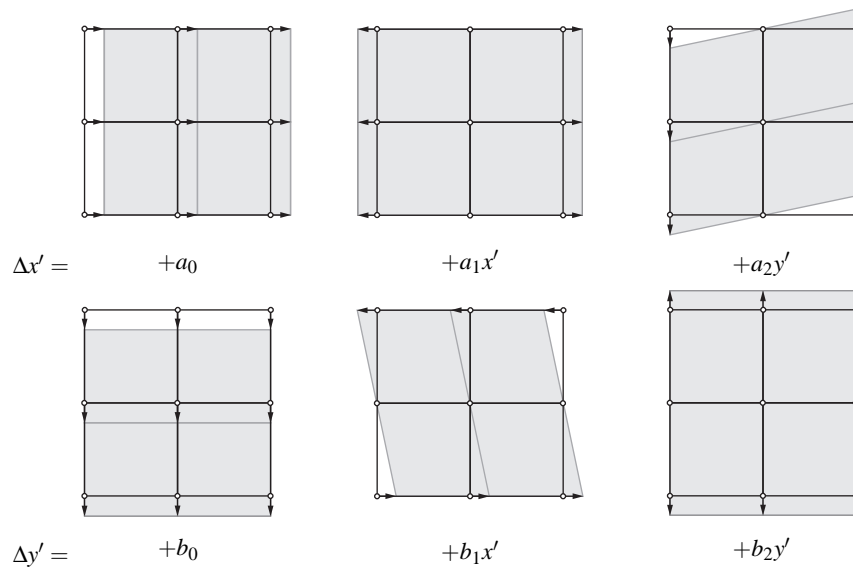
### 7.6.3 Least Squares Surfaces Matching

In den vorherigen **Abschnitten 7.6.1** und **7.6.2** wurden die Transformationsparameter über die Minimierung von Grauwertdifferenzen zwei- bzw. dreidimensionaler Bilddatensätze bestimmt. Natürlich kann der Minimumansatz nach der **MKQ** prinzipiell beliebig attribuiert werden. Liegen z. B. **2,5D**-Daten einer Oberfläche vor, wird also für den Punkt eines Objekts die Höhenkoordinate als Attribut der beiden Lagekoordinaten gespeichert, können kleine Oberflächenausschnitte über die Minimierung der Summe der Quadrate der Höhendifferenzen einander zugeordnet werden. Maas (2000) bestimmt so die Diskrepanzen zwischen benachbarten Streifen eines Flugzeuglaserscanners auf Laserscannerdaten in einer **TIN**-Struktur (**Abbildung 7.11**). Er schneidet dabei das Muster- und das Suchfenster aus den sich überlappenden Streifen der gescannten Objektoberfläche aus, berechnet die notwendigen Gradienten über die Oberflächennormalen des **TIN** und bestimmte den optimalen Transformationsparametersatz über einen klassischen **MKQ**-Ansatz.

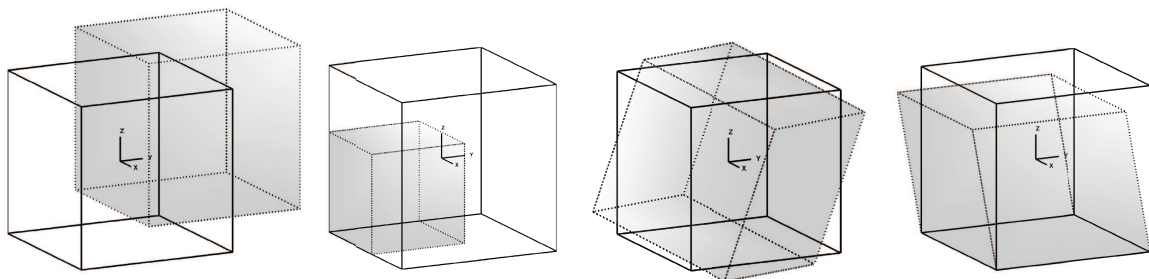
In (Maas 2001) wird ein kombinierter Ansatz zur Bestimmung dreidimensionaler Oberflächenverschiebungen aus Höhen- und Intensitätsdaten von Flugzeuglaserscannern präsentiert. Grün und Akca (2004) stellen einen *Least Squares 3D Surface Matching (LS3D)* genannten Ansatz vor. In Analogie zu (Maas 2000) wird für jeden Punkt einer kleinen Musteroberfläche eine Beobachtungsgleichung aufgestellt und diese im Rahmen einer **MKQ**-Anpassung zur Bestimmung der sieben Parameter einer **3D**-Helmerttransformation durch die Minimierung der Euklidischen Distanzen zwischen zwei oder mehr Oberflächenausschnitten eingeführt. Auch in (Dold & Brenner 2006) wird ein auf **LSM** basierendes Verfahren zur automatischen Registrierung von **3D**-Punktwolken eines terrestrischen Laserscanners präsentiert.

### 7.6.4 Modellerweiterungen

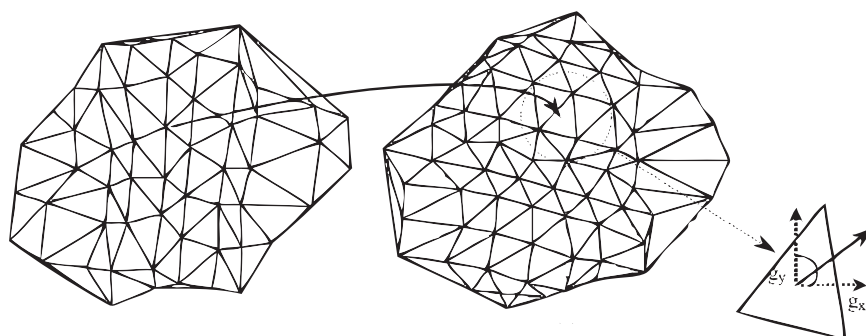
Die flexible Erweiterung des **GMM** zum **GHM** (**Abschnitt 4.4.3**) erlaubt die Berücksichtigung zusätzlicher Nebenbedingungen, was auch in unterschiedlichen **LSM**-Implementierungen genutzt wurde. Bereits Grün (1985) schlägt die simultane Berücksichtigung geometrischer Zwänge, wie z. B. die Forderung nach der Kollinearität von Bildpunkt, Projektionszentrum und Objektpunkt vor. In (D'Apuzzo 2003) werden Zwangsbedingungen zur Einhaltung der Epipolargeometrie eingeführt, um menschliche Körper aus Mehrbildkonfigurationen zuverlässiger rekonstruieren zu können. Der von der *Computer Vision* häufig vermisste Eleganz korrelationsbasierter Verfahren bei zeitlichen Zuordnungsaufgaben (Jähne 2002) kann über die Einführung zusätzlicher Restriktionen nachgekommen werden. Westfeld (2005) entwickelt bspw. einen modifizierten **2D-LSM**-Ansatz, welcher bei der Bestimmung dichter Bewegungsvektorfelder eines Gletschers eine gewisse Kontinuität in der



**Abb. 7.9:** Parameter einer 2D-Affintransformation und ihre Auswirkungen im Bild (nach Kraus 1996).



**Abb. 7.10:** 3D-Affintransformation eines Kuboids: Translation, Maßstab, Rotation, Scherung (v. l. n. r.).



**Abb. 7.11:** Oberflächenausschnitte zweier Laserscannerstreifen in einer TIN-Struktur (Maas 2000).

Gletscherbewegung, folglich also auch in Analogie zur Kontinuitätsgleichung (7.9) im abgebildeten Grauwertfluss fordert. Der LSM-Ansatz mit einem eigentlich diskreten Zeitschritt wurde so als kontinuierliche Bewegung im Raum formuliert. Darüber hinaus legt Westfeld (2005) zugrunde, dass benachbarte Punkte der Gletscheroberfläche ähnliche Bewegungseigenschaften aufweisen, und setzt dies in der Aufstellung räumlicher Restriktionen um. Westfeld und Maas (2010) fordern die Inkompressibilität des extrahierten 3D-LSM-Kuboids bei der volumetrischen Rekonstruktion von durch kleinste Partikel visualisierten gasförmigen oder flüssigen Strömungen.

## 7.7 Diskussion

Das Kapitel bot eine Übersicht zu den für die vorliegende Arbeit relevanten Bereiche der Bewegungsanalyse. Zunächst wurde die Visualisierung einer stattgefundenen Bewegung in der Bildebene besprochen und anschließend mathematisch hergeleitet. Der thematische Schwerpunkt lag bei der Darstellung von Aspekten der digitalen Bildzuordnung. Hier wurden insbesondere grundlegende Prinzipien und Probleme behandelt sowie wichtige Verfahren der Bildkorrelation und Bildzuordnung teils detailliert vorgestellt. Dies ist zum einen wichtig, um die in den Kapiteln 12 und 13 gelisteten Konzepte zur Bewegungserfassung und Bewegungsanalyse zu verstehen, zum anderen basiert die im Kapitel 14 entwickelte und im Kapitel 15 in anwendungsspezifischen Versuchen erprobte Zuordnungsroutine für 3D-Kamerabildaten auf einem integrierten LSM-Ansatz.



# III

Sensormodellierung



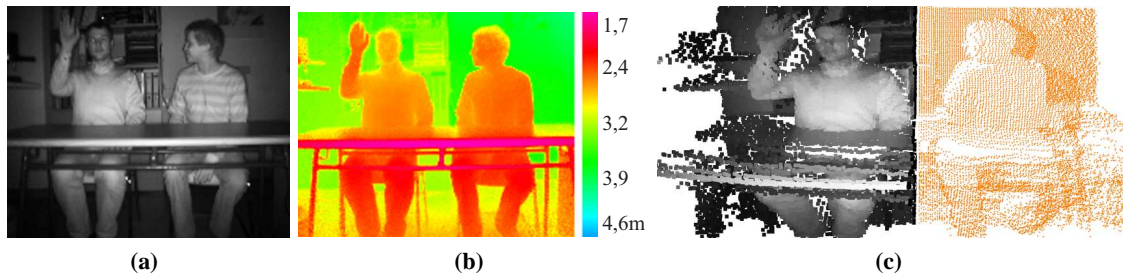
## Distanzmessende Kamera

Distanzmessende Kameras (3D-Kamera<sup>4</sup>, *Range Imaging (RIM) Camera*) auf der Grundlage von *Photomischdetektoren (PMD)* sind Digitalkameras; welche zusätzlich zu einem Grauwertbild synchron ein Entfernungsbild aufnehmen (z. B. Schwarte u. a. 1999; Abbildungen 8.1a und 8.1b). Ein moduliertes Lichtsignal eines Leuchtdiodenbereichs wird von der Objektoberfläche reflektiert und pixelweise durch eine Ladungsträgerschaukel entfernungsselektiv getrennt. Jedes Pixel kann somit als elektrooptischer Entfernungsmesser nach dem Prinzip des Phasenvergleichsverfahrens angesehen werden (Abschnitte 6.3.2 und 8.2). Simultan mit dem Entfernungsbild wird ein aus der Amplitude des reflektierten Signals abgeleitetes Grauwertbild erzeugt. Als Ergebnis liegt eine raumzeitlich aufgelöste 3D-Darstellung des Objektraums vor (Abbildung 8.1c). Das hierfür notwendige geometrische Modell zur Berechnung der 3D-Punktwolken wird in Abschnitt 8.3 eingeführt.

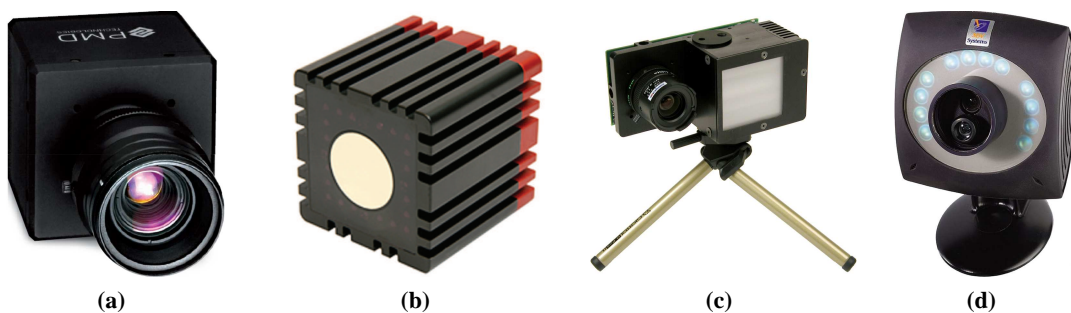
3D-Kameras in der heute bekannten Form entstanden aus einer durch die Schweizer Eidgenossenschaft in Auftrag gegebene Kooperation zwischen dem Forschungs- und Entwicklungszentrum Centre Suisse d'Electronique et de Microtechnique SA (CSEM) und dem Optosensorhersteller CEDES AG (Lange u. a. 1999). Den Markt dominieren derzeit zwei große *spin-off*-Unternehmen: Die PMDTechnologies GmbH aus Siegen (D) als Tochter des ZESS (Zentrum für Sensorsysteme, Universität Siegen) und die MESA Imaging AG aus Zürich (CH), ein Ableger des CSEM. PMD-Tec bietet verschiedene Sensoren nach dem *PMD*-Prinzip an. Das aktuelle Modell *PMD[vision]*<sup>®</sup> CamCube 3.0 (Abbildung 8.2a) weist ein Sensorformat von  $204^2$  px, eine Bildrate von 15 Hz, einen standardmäßigen Eindeutigkeitsbereich von 7,0 m und eine Präzession bei der Distanzmessgenauigkeit von 3 mm bei einer Entfernung von vier Metern auf (PMDTec 2010). Auf demselben Prinzip und ähnlichen Spezifikationen basiert der aktuelle *SwissRanger*<sup>™</sup> SR4000 von MESA (Abbildung 8.2b). Er hat ein Sensorformat von  $(176 \times 144)$  px, eine maximale Bildrate von 50 Hz und eine absolute Messgenauigkeit von einem Prozent der Distanz (MESA 2010). Es existieren weitere Hersteller von *RIM*-Sensorsystemen: Die Firma Canesta, Inc. (Sunnyvale, USA) mit ihrem *EP DevKit*<sup>™</sup> DP200 (Abbildung 8.2c; Canesta 2010) und die von der Microsoft Cooperation (Redmond, USA) aufgekaufte israelische Firma 3DV Systems Ltd. mit der für Spielkonsolenanwendungen industrialisierten *ZCam* (Abbildung 8.2d; Yahav u. a. 2007) seien der Vollständigkeit halber ebenfalls genannt.

Darüber hinaus wird in (Niclass u. a. 2005) ein auf hochempfindlichen und schnellen Photodioden basierender *CMOS*-Bildensor (*Complementary Metal Oxide Semiconductor*) vorgestellt. Im Gegensatz zu der indirekten Laufzeitmessung durch Phasenverschiebung ermöglicht die Implementierung von *Single Photon Avalanche Dioden (SPAD)* eine direkte Laufzeitmessung nach dem Prinzip des

<sup>4</sup> Als 3D-Kameras werden hier Kameras bezeichnet, welche als Einzelsensor simultan aufgenommene Tiefenbilder generieren.



**Abb. 8.1:** Daten einer 3D-Kamera: (a) Grauwertbild. (b) Farbkodiertes Tiefenbild. (c) 3D-Punktwolke mit halbseitiger Texturüberlagerung.

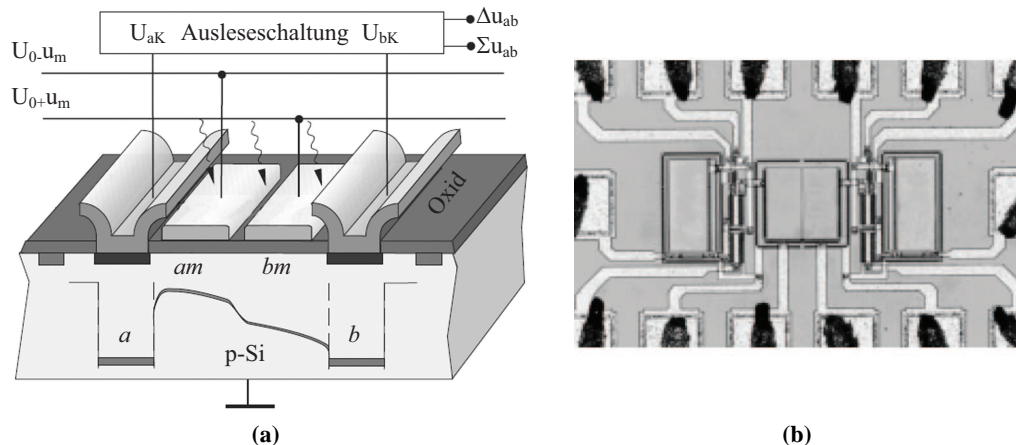


**Abb. 8.2:** Aktuelle 3D-Kameramodelle: (a) PMD[vision]<sup>®</sup> CamCube 3.0, PMDTechnologies GmbH (PMD-Tec 2010). (b) SwissRanger<sup>™</sup> SR4000, MESA Imaging AG (MESA 2010). (c) EP DevKit<sup>™</sup> DP200, Canesta, Inc. (Canesta 2010). (d) Z-Sense<sup>™</sup>, 3DV Systems Ltd. (Yahav u. a. 2007).

Impulsmessverfahrens (Abschnitt 6.3.1).

Die Datenakquisition innerhalb der vorliegenden Arbeit wurde mit den Vorgängermodellen PMD[vision]<sup>®</sup> CamCube 2.0 und SwissRanger<sup>™</sup> SR3k durchgeführt. Die Komponenten eines 3D-Kamerasystems werden in Abschnitt 8.1 vorgestellt. Sofern relevant, wird in den folgenden Ausarbeitungen an entsprechender Stelle auf den jeweils verwendeten Kameratyp hingewiesen, andernfalls sind die gemachten Angaben als repräsentativ für beide 3D-Kameratypen zu verstehen.

Bedeutende Vorteile einer 3D-Kamera im Vergleich zu terrestrischen Laserscannern (TLS) sind vor allem in der Handlichkeit und den geringen Anschaffungskosten zu sehen. Ihr entscheidender Vorteil gegenüber stereophotogrammetrischen Systemen liegt im Wegfall des Arbeitsschritts der stereoskopischen Zuordnung. Eine 3D-Kamera ist somit als monosensorielles System zur Echtzeit-Tiefenbildakquisition mit hoher zeitlicher Auflösung besonders für die in der Arbeit behandelten Anwendungen der menschlichen Bewegungsanalyse geeignet. Darüber hinaus existieren für den Einsatz von 3D-Kameras bereits zahlreiche Ideen und Ansätze in unterschiedlichen Bereichen, welche im Abschnitt 8.4 aufgeführt werden. Ein Vergleich mit konkurrierenden Sensorsystemen erfolgt im Abschnitt 8.5 dieses Kapitels.



**Abb. 8.3:** Der PMD-Halbleiterchips als Kernstück einer 3D-Kamera (nach Xu u. a. 2005): (a) Schematische Darstellung des Querschnitts. Deutlich erkennbar sind die Photogates  $am$  und  $bm$ , welche die aktive Zone definieren sowie die zugehörigen Auslekathoden  $a$  und  $b$ . (b) Mikroskopisches Foto eines einzelnen PMD-Pixel.

## 8.1 Systemkomponenten

Eine 3D-Kamera besteht im Wesentlichen aus den folgenden und im Weiteren kurz beschriebenen Komponenten: Der Bildsensor ist ein neuartiger **Photomischdetektor**-Halbleiter, welcher durch die Detektion und Demodulation eines Lichtsignals gleichzeitig Grauwert- und Entfernungsdaten aufnehmen kann (Abschnitt 8.1.1). Die hierfür notwendige Strahlung wird aktiv von einer Beleuchtungseinheit modelliert und emittiert (Abschnitt 8.1.2) sowie nach einer möglichst diffusen Reflexion von der Kameraoptik wieder aufgefangen (Abschnitt 8.1.3). Schließlich werden die übrigen elektronischen Steuerungselemente im Abschnitt 8.1.4 zusammengefasst.

### 8.1.1 Photomischdetektor

3D-Kameras arbeiten nach dem Prinzip des Phasenvergleichs (Abschnitte 6.3.2 und 8.2). Das dabei zum Einsatz kommende elektronische Verfahren wird als Demodulation oder auch Mischung (Signalmultiplikation) bezeichnet und findet bspw. bei Lock-in-Verstärkern Verwendung (Heckenkamp 2008). Seit den Patentanmeldungen von Spirig u. a. (1995) und Schwarte (1998) wurde an der Entwicklung eines lichtempfindlichen Flächensensors gearbeitet, der dieses Konzept aufgreift und einen phasenempfindlichen Detektor in jeden (Lock-in-)Pixel integriert. Das Ergebnis ist ein **Photomischdetektor** (PMD), der in der Lage ist, ein intensitätsmoduliertes Strahlungsfeld gleichzeitig und ortsabhängig zu detektieren und zu demodulieren (Spirig u. a. 1995; Schwarte 1996; Miyagawa & Kanade 1997; Heinol & Schwarte 1999). Die Abbildung 8.3 zeigt die Topologie eines PMD-Sensors, wie er derzeit in allen gängigen 3D-Kameras verbaut wird.

PMD-Sensoren werden in CMOS-Technologie als **aktive Pixelsensoren** (APS) gefertigt. Diese

monolithische CMOS-Struktur erlaubt die Realisierung von Photosensoren ( $am, bm$ ), also lichtempfindlichen Elementen, und komplexen Ausleseschaltungen ( $a, b$ ) für jedes Pixel als integrierte Detektoreinheit auf einem Chip (Abbildung 8.3). Im Halbleiterbereich unter den Detektorpixeln entstehen, ähnlich dem CCD-Prinzip (*Charge-Coupled Device*), Potentialmulden zur Umsetzung der einfallenden Photonen in Elektronen. Die Potentialmulden vermischen sich durch das Anlegen einer Spannung  $U$  und liefern das gewünschte Korrelationssignal zur Bestimmung der Phasenverschiebung. PMD-Sensoren vereinen somit die Vorteile beider Technologien, CCD und CMOS (Blanc 2001).

Darüber hinaus können weitere Funktionen *on-chip* bzw. pro Pixel integriert werden. Ein wichtiges Beispiel ist die aktive Unterdrückung von unkorreliertem Fremd- bzw. Störlicht (Hintergrundstrahlung; Abschnitt 9.2). Bei Außeneinsätzen verursacht z. B. die Sonneneinstrahlung ein ungünstiges Signal-Rausch-Verhältnis (SNR) und eine vorzeitige Sättigung des Sensors. Eine chipinterne Schaltung zur Fremdlichtunterdrückung (*Suppression of Background Illumination, SBI*) kann diese Effekte durch eine Trennung von korreliertem und nicht korreliertem Lichtanteil in den beiden Potentialmulden deutlich verringern (Möller u. a. 2005). Zusätzlich unterdrückt die integrierte SBI-Schaltung beim Modell PMD[vision]<sup>®</sup> CamCube 2.0 durch thermische Effekte generierte und mit dem eigentlichen Messsignal unkorrelierte Ladungsträger (Ringbeck u. a. 2007).

Der oben beschriebene komplexe Aufbau eines PMD-Elementes resultiert in vergleichsweise großen Pixelkantenlängen von z. B. 45  $\mu\text{m}$  beim aktuellen Modell PMD[vision]<sup>®</sup> CamCube 3.0. Die Sensorgröße und damit die räumliche Auflösung ist mit derzeit 204<sup>2</sup> px als relativ gering anzusehen. Eine Steigerung ist im Zuge der Weiterentwicklung zu erwarten.

Die zeitliche Auflösung des Sensorsystems wird primär über die Integrationszeit (Abschnitt 8.2) und sekundär durch die eingesetzte Soft- und Hardwarelösung festgelegt. Die Hersteller geben hierfür maximal 50 Tiefenbilder pro Sekunde an (MESA 2010).

In den Dissertationsschriften (Lange 2000; Heinol 2001; Justen 2001; Kahlmann 2007) finden sich darüber hinaus gelungene Zusammenfassungen des PMD-Konzepts.

### 8.1.2 Beleuchtungseinheit

Eine ToF-Kamera (*Time-of-Flight*) sendet aktiv und kontinuierlich ein intensitätsmoduliertes Lichtvolumen aus. Die Strahlungsenergie ist dabei abhängig von der Wellenlänge und erhöht sich in jedem Strahlungsquantum mit kürzeren Wellenlängen. Eine hohe Strahlungsenergie ist prinzipiell wünschenswert, besonders im Hinblick auf das maximal erfassbare Messvolumen, jedoch durch Sicherheitsvorschriften limitiert (Abschnitt 8.2). 3D-Kameras arbeiten im Bereich des für das menschliche Auge unsichtbaren nahen Infrarots (NIR) der Laserklasse I zwischen 700 nm bis 1400 nm. Die Produktfamilie SwissRanger<sup>™</sup> arbeitet mit 850 nm, die PMD[vision]<sup>®</sup> CamCube Modelle mit 870 nm (MESA 2010; PMDTec 2010). Auf diese Trägerwelle wird ein charakteristisch bekanntes periodisches Messsignal moduliert, wobei dessen Frequenz den Eindeutigkeitsbereich der Entfernungsmessung festlegt. Typischerweise liegt sie für die aktuell verfügbaren 3D-Kameras zwischen 20 MHz bis 30 MHz, was einer maximal eindeutig auflösbaren Entfernung von 5 m bis 7,5 m entspricht (Ab-

schnitt 8.2).

Die Strahlungsleistung wird durch die Anzahl der **lichtemittierenden Dioden (LED)** bestimmt und beträgt für 3D-Kameras ca. 1 W. Eine hohe Ausstrahlungsleistung kann zwar nicht den Messbereich vergrößern, wohl aber das **SNR** innerhalb des Messvolumens verbessern. Die Messgenauigkeit hängt unter anderem von der Lichtmenge ab, die am beleuchteten Objekt reflektiert wird (vgl. auch **Abschnitt 9.7**).

Die notwendigen **LED** sind raster- oder ringförmig links und rechts neben der bzw. um die gesamte Optik angebracht (**Abbildungen 8.2** und **8.4b**). Besondere Beachtung gilt ihrer symmetrischen Anordnung lotrecht zum Kamerahauptpunkt, infolgedessen sich die durch den Abstand zur optischen Achse induzierten Phasenverschiebungen in ihrer Summe wieder aufheben (**Kahlmann 2007**).

Die spektrale Bandbreite des Transmitters ist klein – beim SwissRanger<sup>TM</sup> SR3k beträgt sie 35 nm (**MESA 2010**) – und gewährleistet die Minimierung des Einflusses fremder Lichtquellen (**Abschnitt 9.2**). Zudem sollten Temperaturschwankungen innerhalb der Beleuchtungseinheit durch adäquate Kühlung vermieden werden, da auch diese die Messgrößen beeinflussen (**Abschnitt 9.1**).

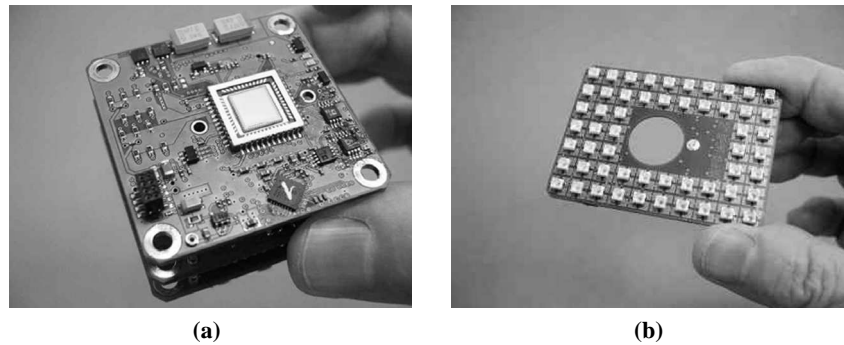
### 8.1.3 Optik

Das optische System einer Kamera sammelt das von der Umgebung reflektierte Licht und bildet es auf dem Sensor ab. Die Abbildungsgeometrie von 3D-Kameras wird im **Abschnitt 8.3** vorgestellt, wobei bereits an dieser Stelle vorweggenommen werden kann, dass sie der Zentralprojektion klassischer 2D-Kameras folgt. Somit unterscheiden sich die verwendeten Objektive nicht von denen herkömmlicher. Als wichtige Eigenschaften sind hier vor allem die große Lichtstärke, die geringen Abschattungseffekte und die Abdeckung eines großen Sichtfelds durch den Einbau von zumeist weitwinkligen Objektiven mit kleiner und fester Blendenzahl sowie die minimale im Rahmen einer geometrischen Kalibrierung zu bestimmenden Objektivverzeichnungen zu nennen. Weiterhin ist ein Bandpassfilter vor der Optik aller Tiefenbildsensoren eingeschaltet, um störendes Hintergrundlicht bestmöglich zu eliminieren. Für den SwissRanger<sup>TM</sup> SR3k beträgt die Bandbreite des optischen Filters bspw. 80 nm (**MESA 2010**).

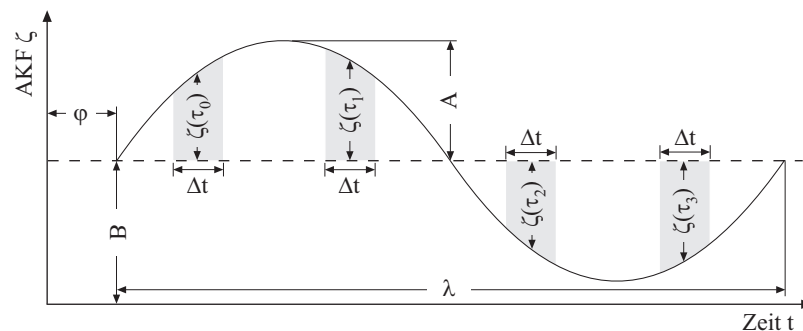
### 8.1.4 Elektronische Steuerungselemente

Die Kameraelektronik einer SwissRanger<sup>TM</sup> SR3k setzt sich modular aus drei Leiterplatten zusammen (**Abbildung 8.4a**). Die oberste enthält den Halbleiter selbst. Sie steuert, überwacht und versorgt ihn mit Strom. Außerdem ist dieses Sensormodul für die Analog-Digital-Wandlung der Halbleiterausgangssignale zuständig. Im Anschluss folgt der Transfer zum Bildverarbeitungsmodul, direkt unterhalb des Sensormoduls. Es verarbeitet die digitalen Rohdaten der vier sequentiellen Intensitätsmessungen und berechnet daraus ein Amplituden- und ein Entfernungsbild (**Abschnitt 8.2**). Zusätzlich können hier erste anwendungsspezifische Bildverarbeitungsfunktionen implementiert werden. Die prozessierten Daten werden zum Datenübertragungsmodul transferiert und per **USB-** oder Netzwerk-schnittstelle an einen angeschlossenen Rechner übertragen.





**Abb. 8.4:** Systemkomponenten eines SwissRanger™ SR3k (Oggier u. a. 2005): (a) Elektronische Leiterplatten. (b) Beleuchtungseinheit.



**Abb. 8.5:** Analyse der AKF mittels Phasenvergleich (nach Kahlmann 2007).

Der beschriebene Aufbau ist repräsentativ für andere 3D-Kameramodelle.

## 8.2 Messprinzip

Distanzmessende Kameras arbeiten nach dem Prinzip des Phasenvergleichs (Abschnitt 6.3.2). Mit einer gerätespezifischen Wellenlänge von ca. 870nm senden sie kontinuierlich NIR aus (Abschnitt 8.1.2). Auf diese Trägerwelle wird ein sinusförmiges Messsignal mit einer Frequenz  $f$  von bspw. 20MHz aufmoduliert. Die ausgesendete Wellenfront wird an Objektoberflächen innerhalb des Gesichtsfelds der Kamera reflektiert und bei diffuser Reflexion teilweise zur Sende-/Empfangseinheit zurückgestrahlt. Das reflektierte Signal trifft aufgrund der zurückgelegten Wegstrecke im Vergleich zu dem ausgesendeten Signal phasenverschoben am Empfänger ein.

Zur Bestimmung der Phasenverschiebung wird das empfangene Signal durch die Abtastung der Autokorrelationsfunktion  $\zeta$  (AKF) rekonstruiert (Abbildung 6.10). Für die Berechnung der drei Parameter Phase  $\varphi$ , Amplitude  $A$  und Offset  $B$  sind mindestens drei Abtastpunkte pro Periode notwendig. In der Praxis erfolgt die Demodulation zu Gunsten einer geringeren Störanfälligkeit in der Regel durch vier sequentielle Intensitätsmessungen  $\zeta(\tau_i)$  in einem Phasenabstand von  $90^\circ$ .

Nach (Lange 2000) ergibt sich die Phase  $\varphi$  nach der Anwendung einer diskreten Fouriertransfor-



mation (DFT) zu:

$$\varphi = \arctan \left[ \frac{\zeta(\tau_1) - \zeta(\tau_3)}{\zeta(\tau_0) - \zeta(\tau_2)} \right] \quad (8.1)$$

Die Rekonstruktion der zur zurückgelegten halben Wegstrecke  $D$  direkt proportionalen Phasenlage  $\varphi$  erfolgt bei 3D-Kameras *on-chip* durch die einzelnen Pixel der CMOS/CCD-Matrix. Jedes (Demodulations-)Pixel enthält also ein Phasenmesssystem, um die eintreffende Welle zu demodulieren und daraus die Distanz  $D$  zu berechnen (vgl. auch [Abschnitt 8.1.1](#)). Aus Gleichung (6.3) ergibt sich:

$$D = \frac{\varphi}{2\pi} \cdot \frac{\lambda}{2} \quad (8.2)$$

mit

$$D_{max} \leq \lambda/2$$

Der maximale Messbereich  $D_{max}$  ist abhängig von der ausgestrahlten Energie der Dioden und, da distanzmessende Kameras mit nur einer Modulationsfrequenz arbeiten, auch abhängig von der Modulationsfrequenz  $f$  des Messsignals ([Abschnitt 8.1.2](#)). In [Abschnitt 6.3.2](#) wurde bereits aufgezeigt, dass die Distanz  $D$  kleiner oder gleich der halben Modulationswellenlänge  $\lambda$  sein muss, um Mehrdeutigkeiten zu vermeiden. Eine Modulationsfrequenz von z. B.  $f = 20$  MHz entspricht laut Gleichung (6.1) einer Modulationswellenlänge von ca.  $\lambda = 15$  m und somit einem Eindeutigkeitsbereich von  $D_{max} \leq 7,5$  m. In der Literatur sind hierzu Ansätze der Arbeitsgruppen um [Jutzi \(2009\)](#) und [Droeschel u. a. \(2010\)](#) zu finden. Bei diesen werden die Wahrscheinlichkeiten von Phasensprüngen unter der Berücksichtigung von Diskontinuitäten im Distanzgradientenverlauf berechnet und so Mehrdeutigkeiten in großdimensionierten Umgebungen ohne die zusätzliche Verwendung von Feinmaßstäben aufgelöst.

Zusätzlich zur Phasenlage kann die Amplitude  $A$  als Maß für die Stärke des empfangenen Signals sowie zur Generierung eines Amplituden- oder auch Grauwertbilds berechnet werden:

$$A = \frac{\delta}{\Delta t \cdot \sin \delta} \frac{\sqrt{[\zeta(\tau_3) - \zeta(\tau_1)]^2 + [\zeta(\tau_0) - \zeta(\tau_2)]^2}}{2} \quad (8.3)$$

mit

$$\delta = \pi \cdot \Delta t / T$$

$$\Delta t \quad \text{Integrationszeit}$$

$$T = f^{-1} \quad \text{Periodendauer}$$

Die Integrationszeit  $\Delta t$  entspricht dem in [Abbildung 8.5](#) dargestellten Zeitintervall zur sequentiellen Abtastung der AKF. Eine Erhöhung der Integrationszeit bedeutet, dass mehr reflektiertes NIR zur Messung verwendet wird. Dies wirkt sich aufgrund eines verbesserten SNR prinzipiell positiv auf die Distanzmessgenauigkeit aus. Vor allem die Aufnahme von Objekten in großer Entfernung pro-

fitiert von einem hohen  $\Delta t$ , da der Photonenverlust aufgrund der längeren Wegstrecke durch hohe Amplitudenwerte ausgeglichen wird. Im Umkehrschluss limitiert diese Eigenschaft jedoch die minimale Messdistanz, da ein hohes  $\Delta t$  bei geringen Objektdistanzen rasch zu einer Sättigung der Pixel führt. Darüber hinaus ist  $\Delta t$  entsprechend der notwendigen Aufnahmegeschwindigkeit zu wählen. Die Wahl der Integrationszeit muss also für jede Aufnahmesituation bezüglich der Reflektivität und mittleren Entfernung der aufzunehmenden Objekte optimal angepasst werden und ist meist einem Kompromiss aus Messgenauigkeit und Aufnahmegeschwindigkeit.

Der (Intensitäts-)Offset  $B$  des empfangenen Signals entspricht der mittleren reflektierten Strahlung und wird im Vergleich zum ausgesendeten Signal hauptsächlich durch die zusätzliche Hintergrundstrahlung hervorgerufen:

$$B = \frac{\zeta(\tau_0) + \zeta(\tau_1) + \zeta(\tau_2) + \zeta(\tau_3)}{4} \quad (8.4)$$

Als weiterführende Literatur sind (Buxbaum 2002; Oggier u. a. 2004; Büttgen u. a. 2005; Kahlmann u. a. 2006; Kahlmann 2007; Ringbeck u. a. 2007) zu nennen. Darüber hinaus werden Aspekte der physikalischen Limitierung des Verfahrens besonders in (Lange 2000) behandelt.

### 8.3 Geometrisches Modell

Die Abbildungsgeometrie folgt einer Zentralprojektion (Abschnitt 5.1). Eine modulierte Lichtquelle sendet aktiv NIR-Strahlung aus. Diese wird am Objekt diffus reflektiert, teilweise durch das Projektionszentrum  $\mathbf{X}_0$  der 3D-Kamera auf den Sensor projiziert und ortsabhängig registriert.

Die Abbildung eines Objektpunkts in die Bildebene ist eindeutig über die in Abschnitt 5.1.5 abgeleiteten Kollinearitätsgleichungen (5.9) beschrieben und soll an dieser Stelle nicht weiter betrachtet werden:

$$\begin{aligned} x' &= F(X_0, Y_0, Z_0, \omega, \varphi, \kappa, x'_0, c, \Delta x', X, Y, Z) \\ y' &= F(X_0, Y_0, Z_0, \omega, \varphi, \kappa, y'_0, c, \Delta y', X, Y, Z) \end{aligned} \quad (8.5)$$

Anders als bei konventionellen 2D-(Digital-)Kameras ist die Umkehrabbildung, also die Projektion eines Bildpunkts in den Objektraum, ebenfalls eindeutig beschrieben. Eine 3D-Kamera erfasst zusätzlich zur Bildinformation eine Entfernungsinformation in Form der Schrägdistanz  $D$  für jedes Pixel. Der für jeden Objektpunkt individuelle und bis dahin unbekannte Maßstabsfaktor  $m$  kann folglich durch diese Distanzmessung  $D$  ersetzt werden. Die kartesischen 3D-Objektkoordinaten  $\mathbf{X}$  können nun durch polares Anhängen für jeden Bildpunkt über seinen Bildvektor  $\mathbf{x}'$  und die zugehörige Distanzmessung  $D$  nach Gleichung (5.7) berechnet werden (Abbildung 8.6):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + (D + \Delta D) \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \left\| \begin{bmatrix} x' - x'_0 - \Delta x' \\ y' - y'_0 - \Delta y' \\ -c \end{bmatrix} \right\| \quad (8.6)$$

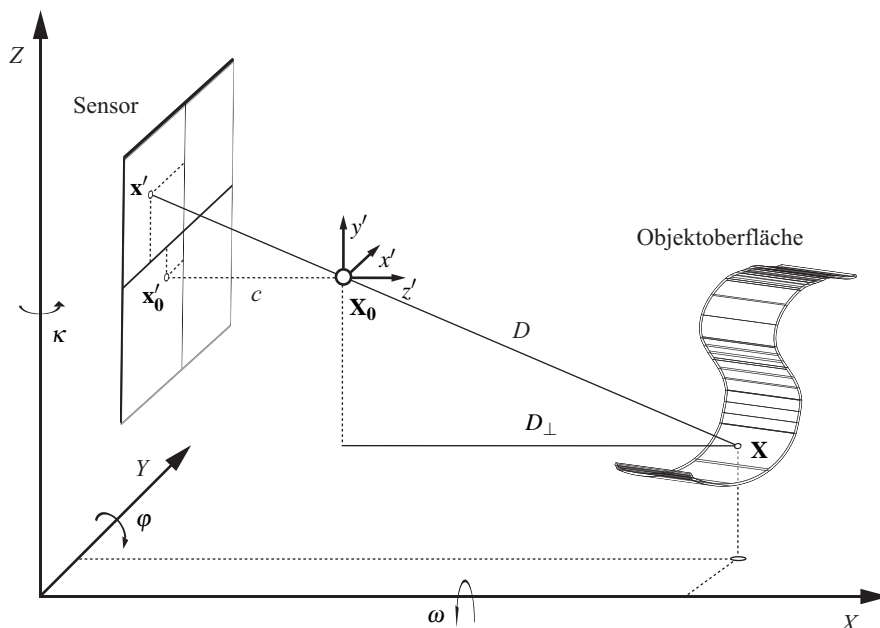


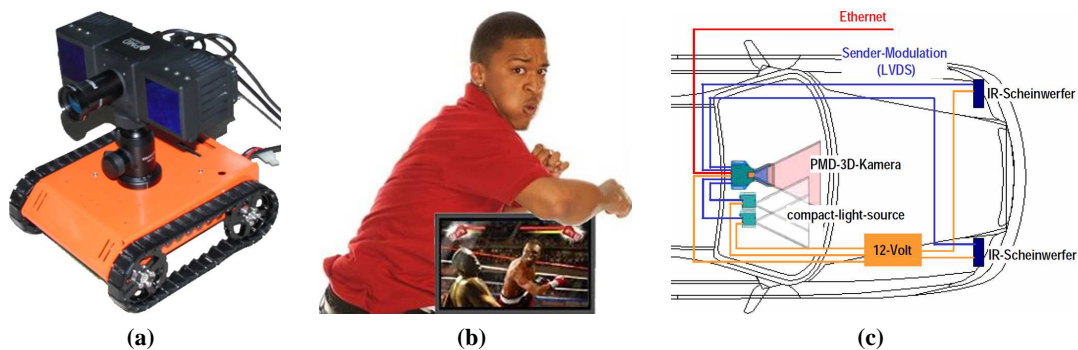
Abb. 8.6: Abbildungsgeometrie einer 3D-Kamera.

In Gleichung (8.6) legt der auf den Hauptpunkt  $\mathbf{x}'_0$  bezogene und um die Korrekturwerte  $\Delta\mathbf{x}'$  der Abbildungsfehler korrigierte Bildvektor  $\mathbf{x}'$  die Richtung zum Objektpunkt fest. Der Objektpunktvektor entspricht somit dem normierten Richtungsvektor  $\|\mathbf{x}'\|$ , skaliert mit der gemessenen Distanz  $D$ . An die Distanzmessung  $D$ , definiert als die Entfernung zwischen dem Projektionszentrum und dem Objektpunkt, wird zusätzlich ein Distanzkorrekturterm  $\Delta D$  angebracht (Kapitel 9 und 10). Falls das Kamerakoordinatensystem  $x'y'z'$  nicht mit einem übergeordneten Objektkoordinatensystem  $XYZ$  zusammenfällt, muss in Gleichung (8.6) die Lage und die Ausrichtung über die Objektkoordinaten des Projektionszentrums  $\mathbf{X}_0$  als Translationsvektor und die Drehmatrix  $\mathbf{R}(\omega, \varphi, \kappa)$  berücksichtigt werden.

Zusätzlich wird häufig die horizontale Distanz  $D_\perp$  benötigt, welche sich direkt aus der Tiefenkomponente des Objektpunktvektors im Kamerakoordinatensystem ableiten lässt (vgl. auch Abbildung 8.6):

$$D_\perp = \|z'\| \quad (8.7)$$

Zur Lösung der Abbildungsfunktion (8.6) ist das einzusetzende Messgerät zu kalibrieren. Neben der Bestimmung der inneren Orientierung und Zusatzparameter ( $c$ ,  $\mathbf{x}'_0$ ,  $\Delta\mathbf{x}'$ ) ist ein weiterer Korrekturterm  $\Delta D$  zur Modellierung von für 3D-Kameras systemspezifischen Fehlern notwendig. Im Kapitel 11 wird eine im Rahmen dieser Arbeit entwickelte integrierte Bündelblockausgleichung vorgestellt, welche alle notwendigen Kalibrierwerte in einem Ansatz simultan berechnet.



**Abb. 8.7:** Anwendungsbeispiele einer 3D-Kamera, I: (a) Autonomer Roboter mit einer montierten PMD[vision]<sup>®</sup> CamCube 2.0 (Schulze 2010). (b) Einsatzmöglichkeiten der ZCam zur Steuerung von Spielen oder Programm-Anwendungen (Pikus 2010). (c) Versuchsträger mit einer A-Muster PMD Kamera für die Anwendungen Fußgängerschutz, *Pre-Crash* und *Stop&Go* (Ringbeck & Hagebecker 2007).

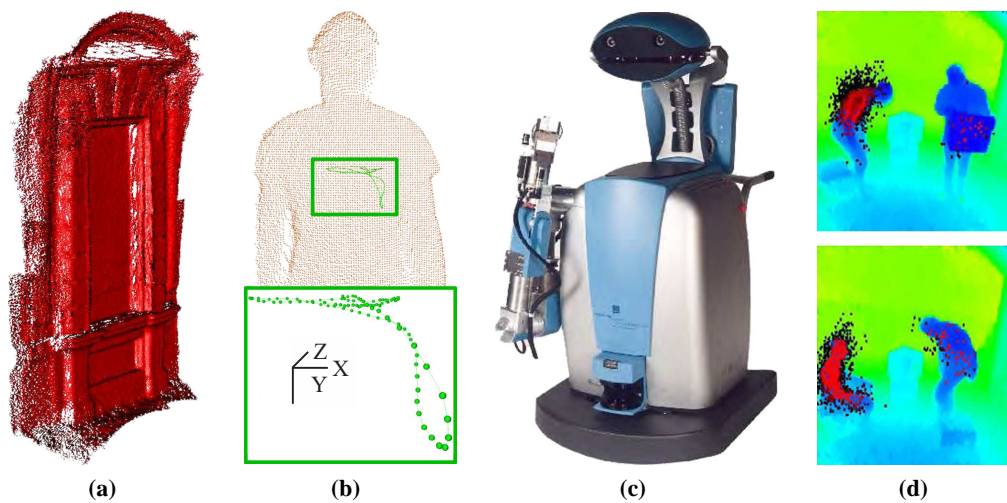
## 8.4 Anwendungsbeispiele

Erste Anwendungen von 3D-Kameras finden sich bspw. in vorausschauenden Sicherheitssystemen in der Automobilindustrie. Göktürk und Tomasi (2004) und Göktürk und Rafi (2005) beschreiben den Einsatz einer 3D-Kamera bei der Kontrolle der Besetzung von Sitzen in einem Fahrzeug und bei der Verfolgung der Köpfe von Fahrzeuginsassen mit dem Ziel der zielgerichteten Auslösung eines Airbags. Zywitza u. a. (2005) und Hsu u. a. (2006) stellen darüber hinaus 3D-Kamera-basierte Assistenzsysteme zur Fahrzeugumfelderfassung für die Applikationen Fußgängerschutz, *Pre-Crash* und *Stop&Go* vor (Abbildung 8.7c).

Ringbeck und Hagebecker (2007) diskutieren den Einsatz von 3D-Kameras in der industriellen Bildverarbeitung für z. B. die vollautomatische Detektion von geometrischen Informationen wie Volumen, Oberflächen oder Füllständen bei Fließbandanwendungen in der Fertigungsindustrie. Frank u. a. (2006) präsentieren neue Formen von Transportwegen in industriellen Arbeitsabläufen, bei welchen 3D-Kameras zur Detektion der Güter verwendet werden können.

Ein wichtiger Anwendungsbereich liegt in der Robotik. Reiser und Kubacki (2007) beschreiben erste Versuche zur Integration einer 3D-Kamera auf dem Endeffektor eines Industrieroboters. Sie können zur Erkennung von Greifobjekten angewendet werden. Kohoutek (2008) demonstriert das Einsatzpotential von 3D-Kameras am Beispiel der Sicherheitsüberwachung des Arbeitsraums von Werkzeugmaschinen und Industrierobotern. Zahlreiche Arbeitsgruppen nutzen bereits eine 3D-Kamera zur Bestimmung der Trajektorie und zur Kartierung der Umgebung eines autonomen Roboters (Weingarten u. a. 2004; Ohno u. a. 2006; Gudmundsson u. a. 2007; May u. a. 2009; Schulze 2010; Abbildung 8.7a). 3D-Kamera-gestützte Rettungsroboter kommen in (Kadous u. a. 2006) und (Sheh u. a. 2006) zum Einsatz.

Das Gebiet der Mensch-Computer-Interaktion (*Human-Computer Interaction*, HCI) befasst sich



**Abb. 8.8:** Anwendungsbeispiele einer 3D-Kamera, II: (a) Fenster des *Castello del Valentino* mit überlagerter 3D-Punktwolke (Chiabrando u. a. 2010). (b) 3D-Punktwolke einer Person mit der 3D-Trajektorie des Körperschwerpunkts (R. Hempel & Westfeld 2009). (c) Haushaltsroboter Care-O-bot (Kubacki & Pfeifer 2005). (d) 3D-Verfolgung von Personen für Überwachungsaufgaben (Kahlmann & Ingensand 2006).

mit der Optimierung von Schnittstellen zwischen Mensch und Maschine. Huber u. a. (2004) zeigen die Echtzeit-Steuerung eines humanoiden Roboters auf der Grundlage von 3D-Kameradaten. Du u. a. (2005) und Yahav u. a. (2007) verwenden RIM-Sensoren als Kamera-Controller zur z. B. Ganzkörpererfassung oder Feinsteuerung durch eine Fingerbewegung bei der Steuerung von Konsolenspielen oder anderen Anwendungen der virtuellen Realität (Abbildung 8.7b). Kubacki und Pfeifer (2005) und Parlitz u. a. (2007) untersuchen die Eignung von 3D-Kameras zur autonomen Navigation und intuitiven Steuerung von Haushaltsrobotern (Abbildung 8.8c). Kohoutek (2009) setzt ToF-Kameras ein, um eine parallele Nutzung von einem Computer durch mehrere Anwender zu ermöglichen.

Verschiedene Arbeiten im Bereich der menschlichen Bewegungsanalyse bewerten die Gestik von Personen aus 3D-Kamerabildsequenzen (Breuer u. a. 2007; Holte & Moeslund 2007). Kahlmann u. a. (2006) weisen auf die schnelle Akquisition von 3D-Daten im Bereich der Gesichtserkennung sowie die Optimierung von Arbeitsabläufen in der Interaktionsumgebung von Mensch und Industrieroboter hin. Kahlmann und Ingensand (2006) und Kahlmann u. a. (2007) detektieren und verfolgen Personen in 3D-Kameradaten für Überwachungsaufgaben (Abbildung 8.8d). Westfeld (2007b) und R. Hempel und Westfeld (2009) bestimmen die Körperorientierung und -distanz zwischen zwei Personen für die Erstellung von Verhaltensstudien (Abbildung 8.8b). Böhme u. a. (2009) entwickeln Algorithmen zur Gesichtserkennung in 3D-Kamerabildsequenzen.

Aufgrund der Schnelligkeit in der Datenakquisition und der geringen Anschaffungskosten diskutieren Westfeld und Maas (2006) den Einsatz von RIM-Sensoren zur Bestandsdokumentation von Industrieanlagen. Chiabrando u. a. (2010) verwenden eine 3D-Kamera zur Erfassung von Gebäuden in der Denkmalpflege (Abbildung 8.8a).

Darüber hinaus dienen ToF-Kameras für die Entwicklung eines neuartigen Echtzeitpositionierungsverfahrens in geschlossenen Räumen (Kohoutek u. a. 2010) und zur geodätischen Vermessung der Topografie des Bettes eines Gebirgsflusses (Kohoutek & Nitsche 2010).

## 8.5 Diskussion

3D-Kameras sind Digitalkameras, bei denen durch die Anwendung von Modulationstechniken zusätzlich zur Bildinformation für jedes Pixel eine Entfernungsinformation vorliegt. Die Kameras nehmen somit simultan Grauwert- und Entfernungsbilder auf. Sie verbinden die Handlichkeit einer Digitalkamera mit dem Potential der dreidimensionalen Datenakquisition eines Laserscanners. 3D-Kameras mit einer Sensorgröße von bis etwa 40.000 px und einer Bildrate von 25 Grauwert- und Entfernungsbildern pro Sekunde werden seit mehreren Jahren angeboten.

Das Kapitel stellte zunächst mehrere 3D-Kameratypen vor und beschrieb die Systemkomponenten, insbesondere den neuartigen und für eine flächenhafte Entfernungsmessung notwendigen PMD-Halbleiter. Das auf einem Phasenvergleich beruhende Messprinzip und das zugrunde liegende Abbildungsmodell wurden, aufbauend auf den im theoretischen Kapitel 5 erarbeiteten Grundlagen, für den 3D-Kamerafall konkretisiert. Zahlreiche im sowohl wissenschaftlichen als auch bereits praktischen Kontext zu findende und im letzten Abschnitt vorgestellte Anwendungsbeispiele aus unterschiedlichsten Bereichen belegen das Potential dieser neuen Technik zur 3D-Oberflächenerfassung in Videorate.

RIM-Kameras stehen in Konkurrenz und Ergänzung zu anderen 3D-Aufnahmeverfahren der Tiefenbildgenerierung wie Laserscanner, Stereokamerasysteme oder Kamera-Projektor-Systeme, die im theoretischen Kapitel 6 bereits näher vorgestellt wurden. Im Folgenden werden einige Vor- und Nachteile vergleichend herausgearbeitet (Tabelle 8.1), welche anwendungsspezifisch zu gewichten sind, in diesem Fall zu Gunsten einer bewegungsanalytischen Aufgabenstellung.

Im Gegensatz zu TLS (Abschnitt 6.3.3) werden Oberflächenmodelle von 3D-Kameras nicht sequentiell, sondern simultan mit hoher Wiederholungsrate aufgenommen. Dies ist bei Anwendungen der Bewegungsanalyse als Vorteil zu werten. Die Punkte des Oberflächenmodells sind bei 3D-Kameras auf einem regelmäßigen Raster im Bildraum angeordnet, was Modellierungsaufgaben und die Übertragung von Algorithmen der Bildverarbeitung deutlich erleichtert (Kapitel 14). Das zusätzlich zum Entfernungsbild aufgenommene Amplitudenbild ist mit dem monochromatischen Bild aus der Intensitätsinformation eines TLS vergleichbar. Die Reichweite von 3D-Kameras ist prinzipbedingt in der Regel auf wenige Meter beschränkt, während TLS Reichweiten von mehr als 1.000 m erreichen. Im Vergleich sind sie auch empfindlicher gegenüber Hintergrundstrahlung. Neben dem Vorteil der Handlichkeit bieten distanzmessende Kameras einen enormen Preisvorteil gegenüber TLS. Technisch bestausgestattete Geräte sind derzeit für etwa € 6.000 erhältlich (ca. 1/5 des Preises eines kostengünstigen TLS), der mittelfristige Marktpreis dürfte bei Produktion in hoher Stückzahl jedoch unter € 1.000 liegen.

Oberflächenmesssysteme (Abschnitt 6.2.1) auf der Basis von Kamera-Projektor-Kombinationen



**Tab. 8.1:** Vergleich von Verfahren zur 3D-Oberflächenerfassung. In Ergänzung zu dem im Fließtext aufgestellten Vergleich konkurrierender 3D-Oberflächenmessverfahren werden hier prägnant die folgenden Kriterien gegenübergestellt: (i) Die Flexibilität, vor allem im Hinblick auf den technischen Aufwand. (ii) Die Komplexität der bis zum Erhalt einer 3D-Information notwendigen Messungen. (iii) Das nach einer Systemkalibrierung zu erwartende Genauigkeitspotential. (iv) Die Möglichkeit, dynamische Vorgänge zu erfassen. (v) Die Systemkosten.

Verfahren	Flexibilität	Komplexität	Genauigkeit	Dynamik	Kosten
Laserscanner	gering	gering	mittel	nein	hoch
Strukturierte Beleuchtung	mittel	mittel	hoch	bedingt	gering
Photometrische Verfahren	gering	mittel	gering	bedingt	gering
Stereoskopische Techniken	gering	hoch	hoch	ja	gering
3D-Kamerasysteme	hoch	gering	gering	ja	mittel

nutzen für die pixelweise Bestimmung von Tiefenkoordinaten Beleuchtungsalgorithmen, welche auf der sequentiellen Projektion mehrerer Codes beruhen. Eine aktive strukturierte Beleuchtung ermöglicht zwar die Messung texturloser Oberflächen im Nahbereich, sie sind damit aber ebenfalls nicht als Echtzeitsystem anzusehen und somit für dynamische Anwendungen nur beschränkt geeignet.

Auch photometrische Oberflächenmessverfahren ([Abschnitt 6.2.3](#)) zeigen ihre Stärken in Regionen mit geringem oder keinem Oberflächenkontrast, was zum Wegfall des sonst zu lösenden Korrespondenzproblems führt ([Abschnitt 7.3](#)). Aufgrund vieler (vereinfachender) Annahmen ist das Genauigkeitspotential begrenzt, eine alternative strenge radiometrische und geometrische Kalibrierung sehr aufwendig sowie die Anwendbarkeit und die Robustheit des Verfahrens im Allgemeinen für photogrammetrische Messaufgaben stark limitiert ([Maas 1997](#)).

Lediglich synchronisierte Stereokamerasysteme ([Abschnitt 6.2.2](#)) liefern simultan aufgenommene Oberflächenmodelle mit hoher Wiederholrate und sind darin mit 3D-Kameras vergleichbar. In der direkten Konkurrenz zwischen den Prinzipien 3D-Kamera und Stereokamerasystem weist die 3D-Kamera den Vorteil des kompakten monosensoriellen Systems auf. Auch ergeben sich durch den Wegfall der Notwendigkeit der stereoskopischen Zuordnung erhebliche Erleichterungen bei der folgenden 3D-Bewegungsanalyse, insbesondere aufgrund der Reduktion von Komplexität und Rechenaufwand. Stereokamerasysteme bieten dagegen derzeit noch eine deutlich höhere Auflösung und ein höheres Genauigkeitspotential. Hierbei ist allerdings zu bedenken, dass sich die Entwicklung von 3D-Kameras noch in den Anfängen befindet und das Potential bei weitem nicht ausgereizt ist.

Neben dem geringen Bildformat und der bereits oben angesprochenen eingeschränkten Reichweite ist die Genauigkeit von durch 3D-Kameras gemessenen Objektpunkten im Bereich kleineren Dezimeter ebenfalls als nachteilig zu betrachten. Im Zuge der Analyse wirkender Fehler, deren anschließenden Modellierung und letztlich der Kalibrierung des gesamten Messsystems kann ein Genauigkeitspotential von wenigen Zentimetern erreicht werden ([Kapitel 9 bis 11](#)). Trotz dieser Ein-

schränkungen stellen distanzmessende Kameras schon heute eine sehr interessante Alternative für Aufgabenstellungen der 3D-Bewegungsanalysen unter kontrollierbaren Beleuchtungsbedingungen dar, bei denen niedrige Sensorkosten und eine hohe zeitliche Auflösung verlangt werden (Teil IV).



## Fehlereinflüsse und deren Korrektur

ToF-Kameras (*Time-of-Flight*) unterliegen einer Vielzahl meist simultan auftretender Fehlereinflüsse. Deren umfassende Untersuchung und anschließende Korrektur im Rahmen verschiedener Kalibrieransätze sind Gegenstand zahlreicher aktueller wissenschaftlicher Arbeiten (z. B. Lange 2000; Luan 2001; Gut 2004; Rapp 2007; Kahlmann 2007; Böhm & Pattinson 2010; Lichti u. a. 2010; Pattinson 2010a; vgl. auch Kapitel 10). Es existieren derzeit noch keine allumfassenden und operationell einsetzbaren Verfahren oder gar kommerziell verfügbare Softwareprogramme zur Bestimmung gerätespezifischer Kalibrierparameter. Für eine zielgerichtete Entwicklung, Implementierung und Validierung von Verfahren zur Auswertung von 3D-Kameradaten sind Kenntnisse über die wirkenden Fehlerquellen, deren Ursprung und Korrekturmöglichkeiten jedoch unabdingbar; die Fehlerbetrachtung und -korrektur ist folglich auch Gegenstand der vorliegenden Arbeit.

Aufgrund des Umfangs der Thematik können im Weiteren nicht alle möglichen Fehlerquellen beschrieben werden. Auf optospezifische Fehler wie z. B. die Abweichungen der realen Abbildungsgeometrie von dem Modell einer zentralperspektivischen Abbildung soll an dieser Stelle nicht eingegangen werden. Distanzmessspezifische Fehler werden insoweit erläutert und in die entwickelte Kalibrieroutine im Kapitel 11 aufgenommen, als ihre Ursachen und Wirkungen ausreichend untersucht sowie in der Literatur beschrieben, sie bestenfalls auch parametrisierbar oder durch entsprechende Aufnahmekonfigurationen vermeidbar und relevant für die vorliegende bewegungsanalytische Aufgabenstellung sind.

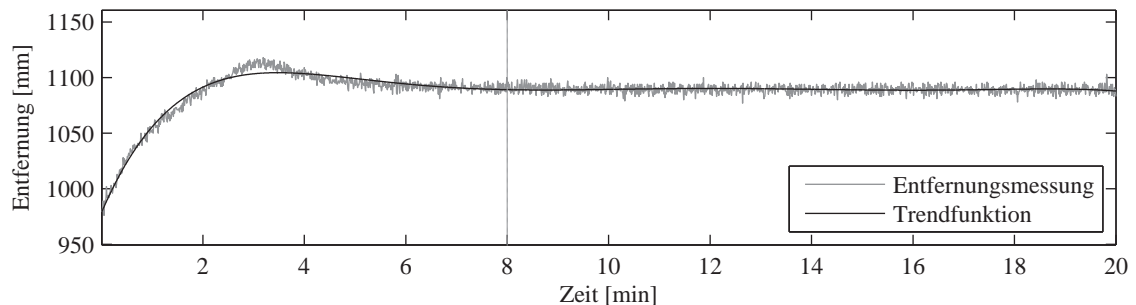
Die Zusammenstellung der folgenden Abschnitte 9.1 bis 9.8 basiert auf eigenen und im Rahmen einer begleitenden Diplomarbeit (M. Hempel 2007) durchgeführten Genauigkeitsuntersuchungen sowie auf Untersuchungsergebnissen fremder Arbeitsgruppen. Sie klassifiziert Fehler, deckt deren Ursachen auf und quantifiziert, inwieweit die eigentliche Messung bzw. die daraus resultierenden Ergebnisse beeinflusst werden. Weiterhin werden Ansätze zur Modellierung und Korrektur aufgezeigt und die Eignung für eine Implementierung in die im Kapitel 11 vorgestellte integrierte Kalibrieroutine bewertet.

### 9.1 Temperatur

Elektronische (Mess-)Geräte unterliegen einem gewissen Einlaufverhalten bis zum Erreichen der internen Betriebstemperatur. Während dieser Zeit kann es zu Schwankungen in den Messergebnissen kommen. Die Abbildung 9.1 zeigt das Einlaufverhalten bei der Distanzmessung einer SwissRanger<sup>TM</sup> SR3k für ein in der Sensormitte gelegenes Pixel bei einer Integrationszeit von 80,8 ms. Im konkreten Fall stabilisieren sich die Ergebnisse der Entfernungsmessungen und entspre-

chend die der zugehörigen Amplitudenmessung nach einer Erwärmungsphase von ca. acht Minuten. Darüber hinaus berichten Karel u. a. (2010) von internen Temperaturschwankungen nach einer Änderung der Integrationszeit im laufenden Betrieb und empfehlen zusätzliche Einlaufzeiten nach jedem Wechsel. Um gerätespezifische Änderungen im Einlaufverhalten zu berücksichtigen, wurde für die im Rahmen dieser Arbeit eingesetzten Geräte eine Einlaufzeit von mindestens 30 Minuten nach der Inbetriebnahme sowie nach der Anpassung der Integrationszeit veranschlagt.

3D-Kamerahersteller weisen eine externe Arbeitstemperatur für den Bereich von ca. 0 °C bis 50 °C aus (MESA 2010; PMDTec 2010). Kahlmann (2007) und Pattinson (2010a) untersuchen in diesem Zusammenhang die Stabilität der gemessenen Distanz verschiedener 3D-Kamerageräte und stellen einen Gradienten von ca. 8 mm/°C fest. Weitere Untersuchungen anderer Arbeitsgruppen sind in der Literatur nicht zu finden und werden auch im Rahmen dieser Arbeit nicht vorgenommen.



**Abb. 9.1:** Das Einlaufverhalten einer 3D-Kamera SwissRanger™ SR3k (Integrationszeit 80, 8ms): Die Entfernungsmessungen unterliegen dem Einfluss interner Temperaturschwankungen und stabilisieren sich nach einer Einlaufzeit von ca. acht Minuten.

## 9.2 Hintergrundstrahlung

Distanzmessende Kameras senden amplitudenmoduliertes **nahes Infrarotlicht (NIR)** aus (Abschnitt 8.1.2). Dieses korrelierte Messsignal vermischt sich mit unkorrelierten Strahlungsanteilen von Fremdlicht, insbesondere dem Sonnenlicht, das in der Regel deutlich stärkere Signalanteile als die aktive Sensorbeleuchtung erzeugt. Die Speicherkapazität der Sensorelemente wird durch Fremdlicht generierte Ladungsträger ausgelastet, im Extremfall kann es zu einer vollständigen Sättigung kommen. Als Konsequenz verringert sich der Speicherbereich des für die Entfernungsmessung relevanten Signalanteils, was folglich eine Verschlechterung des **Signal-Rausch-Verhältnisses (SNR)** und der Messgenauigkeit des gesamten Systems bewirkt (Ringbeck u. a. 2007).

Der Einfluss der Hintergrundstrahlung ist im praktischen Einsatz nicht modellierbar, aber teilweise kompensierbar. 3D-Kamerahersteller wie die PMDTechnologies GmbH kombinieren verschiedene Technologien, um den Einfluss von Fremdlicht zu minimieren: Ein schmalbandiger optischer Filter trennt Signale mit der Wellenlänge des ausgesendeten **NIR** vom restlichen Umgebungslicht (Abschnitt 8.1.3). Zusätzlich verringern eine *Burst*-Überhöhung der Lichtquelle, eine Integrationszeit-

steuerung und eine chipinterne Schaltung zur Fremdlichtunterdrückung (*Suppression of Background Illumination*, SBI; Abschnitt 8.1.1) die Empfindlichkeit des Messsystems gegenüber störender Hintergrundstrahlung (Ringbeck & Hagebecker 2007).

## 9.3 Mehrwegeempfang

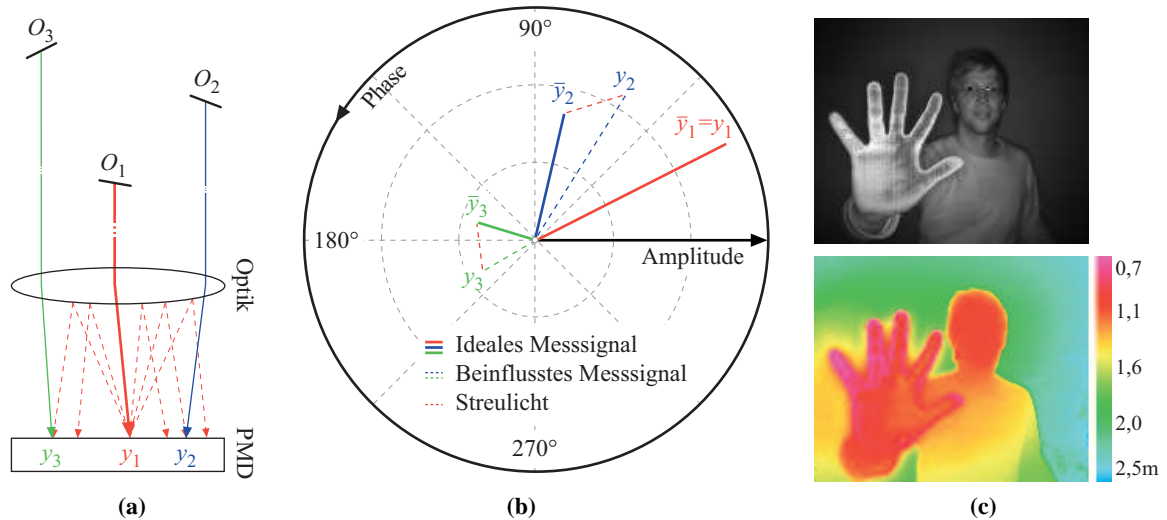
### 9.3.1 Geräteexterne Mehrwegeeffekte

Mehrfachreflexionen im Objektraum werden in der Praxis allgemein als *Multipath* bezeichnet. Die Sensorelemente empfangen Mischsignale aus der direkt reflektierten Strahlung durch das tatsächlich korrespondierende Oberflächenelement, überlagert mit an verschiedenen Objektoberflächen mehrfach reflektierten Messsignalen. Im Vergleich zum direkten Messsignal durchläuft das Mischsignal einen längeren Weg und verursacht einen zu lang registrierten Entfernungswert (z. B. Gudmundsson u. a. 2007; M. Hempel 2007). Für Anwendungen der präzisen GPS-Positionierung (*Globales Positionierungssystem*) existieren unter speziellen Annahmen bereits Ansätze zur Kalibrierung von Trägerphasenmehrwegeeffekten (Wanninger & May 2000; Rost 2011). Ohne Vorwissen über die Geometrie der aufzunehmenden Szene ist eine Modellierung von *Multipath*-Einflüssen bei der Bestimmung von Objektkoordinaten mit einer 3D-Kamera jedoch nicht möglich und lediglich durch adäquate Aufnahmekonfigurationen vermeidbar.

### 9.3.2 Geräteinterne Mehrwegeeffekte

*Scattering* (Streuung) oder auch Linsenreflexion (*lens flare*) sind Begriffe aus der konventionellen Fotografie. Er bezeichnet die aufgrund von Mehrfachreflexionen induzierte Ausbreitung von Teilen des eintreffenden Lichts über den gesamten Sensor. Im Idealfall würde ein Sensorelement lediglich die Strahlung des korrespondierenden Oberflächenelements empfangen. In der Praxis kommt es allerdings innerhalb des optischen Systems und innerhalb des Sensorgehäuses zu Streuungen und Reflexionen, sodass einfallendes Licht mit Streulichtinformationen überlagert wird. Bei einer 3D-Kamera korrelieren die Amplituden- und Distanzmesswerte mit den Messwerten der Nachbarpixel und sind folglich verfälscht (Mure-Dubois & Hügli 2007; Kavli u. a. 2008). Die Auswirkungen auf die Amplitudenwerte sind in der Regel vernachlässigbar, wohingegen die Phasemessungen und die daraus abgeleiteten Distanzwerte vor allem bei 3D-Kamerabildern mit hohen Amplituden und einem deutlichen Tiefenkontrast stark beeinflusst werden können (Karel u. a. 2010; Abbildung 9.2).

Hardwareseitig kann der Einfluss durch den Einsatz spezieller Antireflexionsschichten reduziert werden (Kahlmann & Ingensand 2007b). In der digitalen Bildverarbeitung wird die Wirkung von bandbegrenzten Einflussfaktoren durch eine *Punktspreizfunktion* (PSF, *Point Spread Function*) beschrieben. Mure-Dubois und Hügli (2007) unterstellen eine konstant über den Sensor verlaufende PSF und modellieren den Streulichteinfluss über eine zweidimensionale Gaußsche-Filterfunktion, wobei die getroffenen Annahmen bereits im selben Aufsatz von der Autorenschaft in Frage gestellt



**Abb. 9.2:** Auswirkungen des *Scattering*-Effekts: (a) Aufgrund ihrer unterschiedlichen Entfernung erzeugen die drei Objekte  $O_i$  reflektierte Messsignale  $y_i$  mit unterschiedlichen Amplituden und Phasenwinkeln. Bei Teilen der Rückstrahlung kommt es zu Mehrfachreflexionen zwischen der Optik und dem Sensor (hier am Beispiel für das Objekt  $O_1$ , welches aufgrund der geringen Entfernung das stärkste Echo erzeugt). (b) Das Streulicht von Signal  $y_1$  überlagert die tatsächlichen Signale  $y_{2,3}$ . Unter der Annahme eines perfekt sinusförmigen Messsignals ergeben sich neue überlagerte Signale  $\bar{y}_{2,3}$ . (c) Das Bildpaar visualisiert den Einfluss an realen Messdaten. Die tatsächliche Entfernung der Wand im Hintergrund beträgt ca. 2,0m. Die gemessene Distanz weicht hier vor allem in der Nähe der Hand deutlich ab.

werden. Kavli u. a. (2008) bestimmen lokale PSF für verschiedene Sensorpositionen empirisch. Diese asymmetrisch verlaufenden und nicht-parametrisierten PSF werden iterativ angebracht, ausgehend von Bildregionen mit hohen Amplitudenwerten und reduzieren den Einfluss um bis zu 60%. Dagegen untersuchen Karel u. a. (2010) zunächst umfassend die Auswirkungen von verschiedenen Aufnahmeparametern auf das *Scattering* und behalten sich eine parametrisierte Modellierung für zukünftige Arbeiten vor. Entsprechend kann der Einfluss von Streulicht in der im Kapitel 11 vorgestellten integrierten Kalibrieroutine nicht berücksichtigt werden.

## 9.4 Fixed Pattern Noise

Als *Fixed Pattern Noise* (FPN) wird ein für jedes Pixel unterschiedlicher aber konstanter Messfehler bezeichnet, der die Bilder konventioneller CMOS-Kameras (*Complementary Metal Oxide Semiconductor*) aufgrund ungleicher Materialeigenschaften der Sensorelemente überlagern kann (Lange 2000; Jähne 2002). Diese herstellungsbedingten Unterschiede der einzelnen CMOS-Elemente in z. B. der Dicke der Metalloxide oder der Größe der lichtempfindlichen Fläche bewirken unterschiedliche Empfindlichkeiten (*Offset*) und unterschiedliche Eigenschaften in der Signalverstärkung (*Gain*).

Die Elektronik von PMD-Laufzeitkameras (*Photomischdetektor*) unterliegt dem selben Problem

(Kahlmann u. a. 2007; Fuchs & Hirzinger 2008; May u. a. 2009). In Folge wird individuell für die Distanz eines jeden Pixels ein additiver amplitudenunabhängiger *Offset* und ein multiplikativer amplitudenabhängiger *Gain* angebracht (Hussmann & Edeler 2010). Die Korrekturwerte sind auch hier konstant und werden bei den aktuellen 3D-Kameras vom Hersteller durch entsprechende Korrekturmatriizen (Kahlmann 2007) oder Mehrfachsignalabtastung (Luan 2001) berücksichtigt. Der Einfluss des FPN wird im Weiteren als korrigiert angesehen.

## 9.5 Distanzabhängige Fehlereinflüsse

Für eine qualitative und quantitative Einschätzung der Distanzmessgenauigkeit werden Vergleichsstrecken für ein in der Sensormitte gelegenes Pixel einer 3D-Kamera SwissRanger™ SR3k für unterschiedliche Integrationszeiten auf einer interferometrischen Kalibrieranlage des Geodätischen Instituts der TU Dresden bestimmt. Der resultierende Soll-Ist-Verlauf ist in [Abbildung 9.3](#) (unten) dargestellt. Dessen sukzessive Analyse führt zu einer Aufspaltung in einen linearen (mittig) und einen zyklischen Trend (oben). Die Aufdeckung der fehlerinduzierenden Ursachen und deren Modellierung ist Gegenstand der folgenden zwei [Abschnitte 9.5.1](#) und [9.5.2](#). In (Kahlmann 2007) wird diese Untersuchung auf mehrere über den Sensor verteilte Pixel erweitert.

### 9.5.1 Linearer Fehler

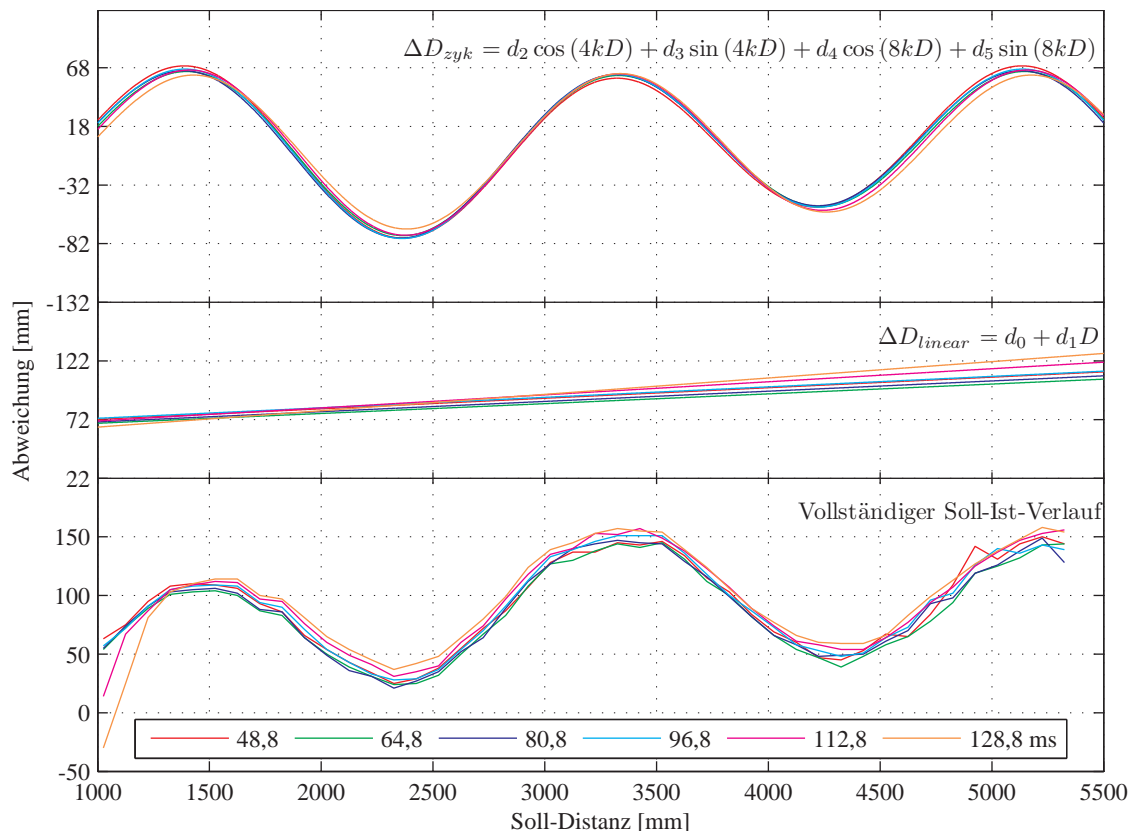
Für die Aufstellung des geometrischen Modells wird in [Abschnitt 8.3](#) die gemessene Schrägdistanz  $D$  als Entfernung zwischen dem Projektionszentrum und dem Objektpunkt angenommen. Etwaige Abweichungen von dieser in der Photogrammetrie üblichen Definition werden durch einen konstanten Verschiebungsparameter  $d_0$  berücksichtigt.

Darüber hinaus ist aus [Abbildung 9.3](#) erkennbar, dass der Fehler mit zunehmender Entfernung wächst. Durch instrumentell bedingte Abweichungen der Modulationsfrequenz  $\lambda$  kann es zu Maßstabsabweichungen kommen (Joeckel & Stober 1999). Diese Frequenzabweichung verursacht eine streckenproportionale Entfernungsabweichung, welche den hier dokumentierten linearen Trend erklären kann. Als eine Ursache für die Frequenzabweichung geben Joeckel und Stober die Temperaturabhängigkeit des Schwingquarzes an, welcher für die Erzeugung der Resonanzfrequenz verantwortlich ist. Neben dem konstanten *Offset*  $d_0$  ergibt sich so ein entfernungsabhängiger Korrekturwert  $d_1$ . Zusammen bilden sie den in [Kapitel 11](#) zu bestimmenden linearen Distanzkorrekturwert  $\Delta D_{lin}$ :

$$\Delta D_{lin} = d_0 + d_1 D \tag{9.1}$$

### 9.5.2 Zyklischer Fehler

In der Modulation des Messsignals kann es zu Interferenzen mit zusätzlichen Oberwellen kommen, also zu Überlagerungen mit ganzzahligen Vielfachen der sinusförmigen Grundschwingung. Die Annahme einer harmonischen sinusförmigen Beleuchtung ist aufgrund der asymmetrischen [Autokorre-](#)



**Abb. 9.3:** Darstellung der Soll-Ist-Abweichungen für das zentrale Sensorpixel einer 3D-Kamera SwissRanger™ SR3k für unterschiedliche und im Plot farbige dargestellte Integrationszeiten (unten). Mittels einfachen Regressionsanalysen lässt sich das Signal in einen linearen Anteil und einen zyklischen Anteil trennen. Der additive Term des linearen Anteils liegt bei ca. 60 mm. Der multiplikative Term  $d_1$  liegt bei ca.  $1e^{-2}$  (mittig), was einer Abweichung von ca. 50 mm bei 5 m Entfernung entspricht. Deutlich erkennbar sind die periodischen Schwankungen im zyklischen Anteil mit Amplituden von etwa  $\pm 70$  mm (oben).

lationsfunktion (AKF) nicht mehr gültig und resultiert in falsch bestimmten Phasenverschiebungen und folglich auch in falschen Distanzwerten. Dieser zyklische Effekt, welcher für 3D-Kameras bereits durch Lange (2000) prädiziert wurde, ist für verschiedene Integrationszeiten eines Sensorelements in [Abbildung 9.3](#) (oben) dargestellt. Da der funktionale Verlauf stetig ist, kann er mithilfe einer *Lookup-Tabelle* (LUT) oder einer geeigneten Kalibrierfunktion hergestellt werden.

In (Kahlmann u. a. 2006) wird eine mehrdimensionale LUT zur Korrektur als ausreichend betrachtet, wohingegen in (Fuchs & Hirzinger 2008; Schiller u. a. 2008; Kim u. a. 2008) Polynome höheren Grades und in (Lindner & Kolb 2006) *B-Splines* zur Modellierung der zyklischen Fehler verwendet werden. Der Fehlereinfluss wird durch diese Ansätze deutlich verringert, ein Rückschluss auf die tatsächliche Fehlerquelle ist jedoch erst durch eine streng geometrisch-physikalische Modellierung möglich (Joeckel & Stober 1999). Pattinson (2010b), Böhm und Pattinson (2010) sowie Lichti u. a. (2010) untersuchen in diesem Zusammenhang die Eignung verschiedener sinusförmiger Funktionen

und belegen, dass die folgenden zwei Wellen den zyklischen Phasenfehler hinreichend genau approximieren:

$$\Delta D_{zyk} = d_2 \cos(4kD) + d_3 \sin(4kD) + d_4 \cos(8kD) + d_5 \sin(8kD) \quad (9.2)$$

mit

$$k = \frac{2\pi}{\lambda} = \frac{2\pi f}{c} \quad \text{Modulationswellenzahl}$$

Die Wellenlängen der sinusförmigen Korrekturfunktionen entsprechen  $1/4$  bzw.  $1/8$  der als konstant angenommenen Modulationswellenlänge  $\lambda$ ; die Verschiebung der Phasen werden durch die Kosinusanteile berücksichtigt. Somit verbleiben einzig die Amplituden  $d_2$  bis  $d_5$  als unbekannte Wellenparameter, welche in das funktionale Modell der im Rahmen dieser Arbeit entwickelten Kalibrierstrategie eingeführt werden (Kapitel 11).

## 9.6 Signalwegeeffektbedingte Fehler

Aufgrund von Verzögerungen bei der Signalausbreitung kommt es, neben dem in Abschnitt 9.4 diskutierten FPN, zu einer weiteren sensorpositionsabhängigen Phasenverschiebung beim Aussenden und Empfangen der NIR-Strahlung (Du u. a. 2005; Fuchs & Hirzinger 2008). Der Einfluss der Sensorposition wird in dieser Arbeit über den Bildradius  $r'$ , also den Abstand der um die Abbildungsfehler  $\Delta x'$  reduzierten Bildkoordinaten  $\mathbf{x}'$  zum Bildhauptpunkt  $\mathbf{x}'_0$ , durch die folgende Korrekturgleichung berücksichtigt:

$$\Delta D_{x'y'} = d_6 r' \quad (9.3)$$

mit

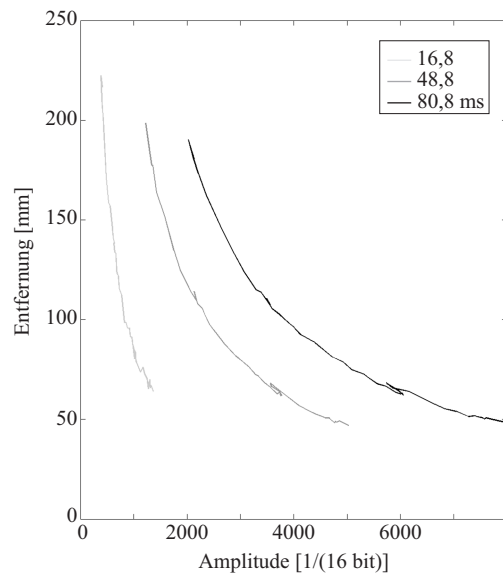
$$r' = \sqrt{(x' - x'_0 - \Delta x')^2 + (y' - y'_0 - \Delta y')^2} \quad \text{Bildradius}$$

## 9.7 Reflektivität

Die in jedem PMD-Element demodulierte Amplitude ist ein Maß für die Stärke des empfangenen Signals und somit gleichzeitig ein Maß für die Reflektivität des korrespondierenden Oberflächenelements (Abschnitt 8.2). Der Reflexionsgrad als Verhältnis zwischen einfallender und reflektierter Strahlung ist hierbei vor allem von der Oberflächenbeschaffenheit des Aufnahmeobjekts, also z. B. von Farbe, Form, Größe und Material abhängig (Lange 2000). Für eine umfassende Kalibrierung ist nun von Interesse, ob sich die Objektreflektivität auch auf die Distanzmessgenauigkeit einer 3D-Kamera auswirkt.

Oprinescu u. a. (2007) führen eine pixelweise amplitudenbasierte Distanzkorrektur für solche Sensorelemente durch, die einen hohen Entfernungswert aufweisen. Nach dem Vorbild radiometrischer





$$\Delta D_{amp} = -\frac{d_7}{d_8}A + \sqrt{\frac{d_7^2}{d_8^2}A^2 - 2\frac{d_9}{d_8}A + \frac{1}{d_8}} \quad (9.4)$$

mit

$A$  Amplitudenwert  $[0, 1]$

$d_7, d_8, d_9$  Hyperbelparameter

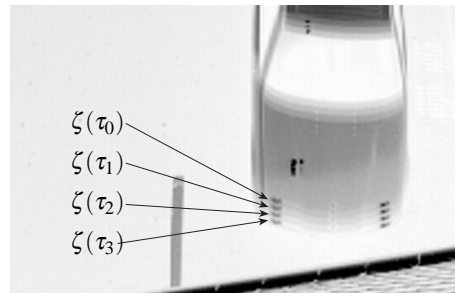
**Abb. 9.4:** Einfluss der Amplitude auf die Entfernungsmessung (Kahlmann 2007): Der Plot zeigt den Zusammenhang zwischen den Amplituden- und Distanzwerten für ein Pixel der 3D-Kamera SwissRanger™ SR3k für unterschiedliche Integrationszeiten. Mit sinkender Reflektivität erhöht sich der gemessene Entfernungswert.

Laserscanner-Korrekturmodelle (Kaasalainen u. a. 2008) schlagen Jaakkola u. a. (2008) eine empirische Kalibrierung der Amplitudenwerte vor. Radmer u. a. (2008) erstellen mehrdimensionale LUT zur Distanzkorrektur in Abhängigkeit von der Integrationszeit und der Amplitude. Lindner und Kolb (2007) approximieren den amplitudenbasierten Fehler mit *B-Splines*. Kahlmann und Ingsand (2007a) berichten von einem signifikanten Einfluss der Amplitude auf die Distanzmessung von mehreren Zentimetern und weisen auf die Möglichkeit einer Korrelation zum *Scattering*-Effekt hin (Abschnitt 9.3). In seiner Dissertationsarbeit zeigt Kahlmann (2007) Ähnlichkeiten des funktionalen Zusammenhangs zwischen den Amplituden-/Entfernungswerten und einer Hyperbel. Karel und Pfeifer (2009) adaptieren das in (Kahlmann 2007) aufgestellte Modell, stellen verbleibende Restsystematiken fest und überführen es schließlich in eine allgemeine Hyperbelform, reduziert um den nicht relevanten Achsoffset-Parameter. Die resultierende Gleichung (9.4) zur Reduktion der Distanzmessung um den Einfluss der Reflektivität eignet sich für eine Integration in das funktionale Modell der im Kapitel 11 vorgestellten integrierten Bündelblockausgleichung.

## 9.8 Bewegungsartefakte

Ein weiterer dem PMD-Aufnahmeprinzip geschuldeter Fehlereinfluss sind Bewegungsartefakte. Sie bilden sich, wenn sich das Objekt und/oder die Kamera zwischen den vier konsekutiven Abtastmessungen  $\zeta(\tau_i)$  bewegt (Abschnitt 8.2) und werden durch schnellere Bewegungen sowie längere Integrationszeiten begünstigt. Es kommt genau dann zu einer Artefaktbildung, wenn während der





**Abb. 9.5:** Visualisierung von Bewegungsartefakten: Das inverse ToF-Amplitudenbild zeigt die Aufnahme eines sich schnell bewegenden Fahrzeugs. Deutlich zu erkennen ist die klar separierte vierfache Artefaktbildung im Bereich der Scheinwerfer zu jedem Abtastzeitpunkt  $\zeta(\tau_{[0,3]})$ .

Demodulation Phasensprünge auftreten. In diesem Zusammenhang unterscheiden [Lindner und Kolb \(2009\)](#) drei artefaktinduzierende Ursachen:

- Laterale Bewegungen verursachen an den Grenzen eines sich bewegenden Objekts Mischsignale aus Vorder- und Hintergrundstrahlung ([Abschnitt 9.3](#)).
- Axiale Bewegungen verursachen zusätzliche Phasenänderungen aufgrund variierender Objektdistanzen.
- Hohe Texturgradienten verursachen ungleiche Phasenwerte aufgrund unterschiedlicher Reflexionseigenschaften ([Abschnitt 9.7](#)).

[Lottner u. a. \(2007\)](#) versuchen diesen Effekt rechentechnisch über eine kantenerhaltende Glättung kombinierter 2D/3D-ToF-Bilder zu kompensieren. [M. O. Schmidt \(2008\)](#) zeigt am Beispiel einer texturbasierenden Artefaktbildung, dass dieser Fehlereinfluss sowohl im Amplituden- als auch im Entfernungsbild auftreten kann und die entstandenen Artefaktbereiche im Gegensatz zur Bewegungsunschärfe konventioneller 2D-Kameras klar separiert sind ([Abbildung 9.2](#)). [Lindner und Kolb \(2009\)](#) schätzen zunächst den optischen Fluss des sich bewegenden Objekts aus dem Amplitudenbild ([Abschnitt 7.5](#)) und passen anschließend das gerechnete Entfernungsbild entsprechend der geschätzten Objektbewegung an. [Hussmann u. a. \(2010\)](#) schlagen eine hardwarebasierte Korrektur der Amplitudenbilder in Echtzeit vor.

Auftretende Bewegungsunschärfe wird bei der verwendeten 3D-Kamera PMD[vision]<sup>®</sup> CamCube 2.0 teilweise hardwareseitig korrigiert (*Suppression of Motion Blur, SMB*; [PMDTec 2010](#)). Inwieweit verbleibende Bewegungsartefakte die Ergebnisse des in [Kapitel 14](#) vorgestellten Ansatzes zur Verfolgung von Objekten in Tiefenbildsequenzen beeinflussen, wird im Rahmen der vorliegenden Arbeit nicht betrachtet und bleibt somit zukünftigen Untersuchungen vorbehalten.

## 9.9 Diskussion

Zahlreiche Arbeitsgruppen als auch die Hersteller von 3D-Kamerageräten setzen sich mit der Beschreibung, Erfassung und Modellierung von auf eine 3D-Kameradistanzmessung wirkenden Feh-

lereinflüssen auseinander. Das Kapitel listete die hauptsächlichsten auf. Es stellte darüber hinaus die aus begleitend durchgeführten Genauigkeitsuntersuchungen resultierenden bzw. die in der aktuellen Literatur zu findenden Korrekturansätze vor und wählte diejenigen aus, für welche ein geometrisch-physikalisch streng parametrisiertes Fehlermodell gilt:

- Die linearen und zyklischen Distanzkorrekturwerte  $\Delta D_{lin}$  und  $\Delta D_{zyk}$  modellieren distanzabhängige Einflussfaktoren.
- Der signalwegeeffektbedingte Fehlerkorrekturterm  $\Delta D_{x'y'}$  kompensiert den Einfluss der Sensorposition.
- Der amplitudenabhängige Entfernungskorrekturterm  $\Delta D_{amp}$  berücksichtigt die unterschiedlichen Reflektivitätseigenschaften von Objekten.

In der Summe ergibt sich folgendes Modell zur Korrektur distanzmessspezifischer Fehlereinflüsse, die, zusätzlich zu den hier nicht diskutierten optospezifischen Fehlern, im Rahmen einer integrierten Bündelblockausgleichung im [Kapitel 11](#) geschätzt werden:

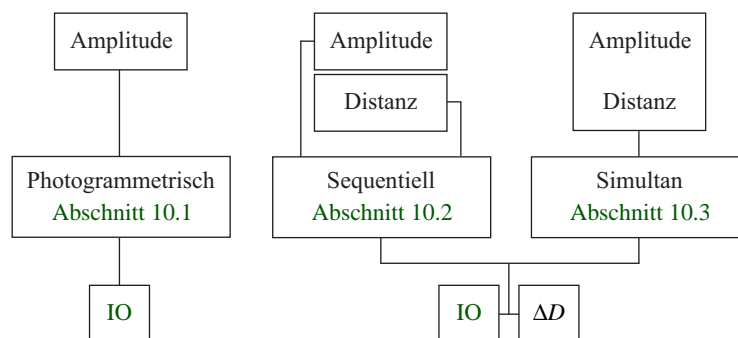
$$\Delta D = \Delta D_{lin} + \Delta D_{zyk} + \Delta D_{x'y'} + \Delta D_{amp} \quad (9.5)$$

Darüber hinaus wurden aufgezeigt, inwieweit die Einflüsse verbleibender Fehlerquellen eliminiert oder zumindest reduziert werden können:

- Die Berücksichtigung des Einlaufverhaltens stabilisiert die Gerätetemperatur und minimiert deren Einfluss auf das Messergebnis.
- Die Hintergrundstrahlung wird seitens des Geräteherstellers durch optische Filter und elektronische Steuerungselemente weitgehend unterdrückt.
- Geräteinterne Mehrwegeeffekte werden hardwareseitig reduziert.
- Geräteextern auftretende Mehrwegeeffekte können nur durch eine entsprechende Aufnahmeconfiguration vermieden werden.
- Die prinzipbedingte Bildung von Bewegungsartefakten wird bei einigen [3D-Kameramodellen](#) chipintern korrigiert.

## Kalibrierstrategien

Im vorangegangenen [Kapitel 9](#) wurde bereits gezeigt, dass sich derzeit zahlreiche Arbeitsgruppen aus unterschiedlichen wissenschaftlichen Disziplinen, wie z. B. der Informatik, der Physik, der Elektrotechnik und nicht zuletzt der Geodäsie und Photogrammetrie, mit Untersuchungen zur Messperformance von Laufzeitkameras beschäftigen. Basierend auf diesen Untersuchungen können entsprechende Fehlermodelle und Kalibrierstrategien zur Modellierung der wirkenden Fehler entwickelt werden. Das folgende Kapitel klassifiziert die in der Literatur veröffentlichten Kalibrieransätze, angefangen bei konventionellen photogrammetrischen Kalibrierungen, über sequentielle Verfahren zur Bestimmung der inneren Orientierung mit Zusatzparametern und der Distanzkorrektur bis hin zur gemeinsamen Modellierung aller auf eine 3D-Kamera wirkender Fehlerquellen in einem integrierten Ansatz ([Abbildung 10.1](#)). Es erläutert deren Prinzipien kurz und nennt die wichtigsten Ergebnisse.



**Abb. 10.1:** Klassifikation von 3D-Kamerakalibrieransätzen hinsichtlich der verwendeten Eingabekanäle (Beobachtungen, gemessen im Amplituden- bzw. Distanzbild) und der resultierenden Kalibrierwerte (Parameter der inneren Orientierung und Zusatzparameter bzw. Distanzkorrekturterme).

### 10.1 Photogrammetrische Kalibrierung

3D-Kameras bilden Objektpunkte nach dem Prinzip der Zentralprojektion auf dem Bildsensor ab ([Abschnitt 8.3](#)). Die Abweichungen von der idealen geometrischen Projektion werden hierbei durch die Parameter der inneren Orientierung und Zusatzparameter abgefangen ([Abschnitt 5.1.3](#)). Deren Bestimmung erfolgt im Rahmen einer Kamerakalibrierung, welche Standard in der digitalen Photogrammetrie ist (z. B. [Fraser 1997](#)).

Das Prinzip der photogrammetrischen Kamerakalibrierung kann vollständig auf 3D-Kameras übertragen werden. Als flexible, einfach zu handhabende und robuste Methode zur Bestimmung der Parameter der inneren Orientierung und Zusatzparameter einer 3D-Kamera bietet sich die Referenzfeld-

kalibrierung an (Abschnitt 5.3.3). In der Literatur finden sich zahlreiche Arbeitsgruppen, welche die innere Orientierung auf diese Weise durch die Messung von mit übergeordneter Genauigkeit bekannten Referenzkoordinaten in den Amplitudenbildern bestimmen (Reulke 2006; Santrac u. a. 2006; Beder & Koch 2008). Bei einer ausreichenden Tiefenausdehnung des Punktfelds und einer hohen Punktzahl genügt in der Regel ein räumlicher Rückwärtsschnitt anhand eines Einzelbilds.

Westfeld (2007a) bestimmt die Kameraparameter im Rahmen einer Bündelblockausgleichung mit Selbstkalibrierung anhand mehrerer formatfüllender Konvergentaufnahmen eines Testfelds, inklusive Rotationsstrategien zur Vermeidung von Korrelationen zwischen den Parametern der inneren und äußeren Orientierung (Abbildung 10.2; Abschnitt 5.4.3). Aus dieser Vorarbeit für die im Kapitel 11 vorgestellte integrierte Bündelblockausgleichung resultieren die in Tabelle 10.1 gelisteten Ergebnisse der photogrammetrischen Kalibrierung einer SwissRanger™ SR3k. Die a-posteriori Standardabweichung  $\hat{\sigma}_0$  der Gewichtseinheit liegt bei  $1/30$  px. Für eine ausführlichere Einschätzung sei an dieser Stelle auf die Ergebnisanalyse der integrierten Bündelblockausgleichung in Abschnitt 11.8 verwiesen.

**Tab. 10.1:** Die Parameter  $x_i$  der inneren Orientierung und deren a-posteriori Standardabweichungen  $\hat{\sigma}_{x_i}$  einer SwissRanger™ SR3k.

	$c$ [mm]	$x'_0$ [mm]	$y'_0$ [mm]	$A_1$	$A_2$	$A_3$	$B_1$	$B_2$	$C_1$	$C_2$
$\hat{x}_i$	7,899	0,191	0,697	$2,50e^{-3}$	$4,86e^{-5}$	0	$1,83e^{-4}$	$1,82e^{-4}$	$3,94e^{-4}$	$8,67e^{-4}$
$\hat{\sigma}_{x_i}$	$5,37e^{-3}$	$1,58e^{-2}$	$1,16e^{-2}$	$1,33e^{-4}$	$9,44e^{-6}$	fix	$6,26e^{-5}$	$5,54e^{-5}$	$1,79e^{-4}$	$1,71e^{-4}$

## 10.2 Sequentielle photogrammetrische und Distanzkalibrierung

Neben der Bestimmung der inneren Aufnahmegeometrie ist bei 3D-Kameras zusätzlich die Kalibrierung der Distanzmessung notwendig (Kapitel 9). Die im Folgenden vorgestellten Kalibrierverfahren arbeiten sequentiell. Während die Kamerageometrie von allen Arbeitsgruppen nach dem in Abschnitt 10.1 vorgestellten Prinzip hergestellt wird, unterscheiden sich die Ansätze zur Berechnung der Distanzkorrekturterme vor allem hinsichtlich der Bestimmung von Referenzstrecken.

Kahlmann und Ingensand (2004; 2006) und Kahlmann u. a. (2006) ermitteln über die definierte Zielverschiebung auf einer interferometrischen Komparatorbank für jedes Pixel einen additiven LUT-Distanzkorrekturwert (Abbildung 10.3a). Die Abweichungen der Distanzmessungen aufgrund des *Fixed Pattern Noise* (FPN) können so von bis zu  $\pm 30$  cm vor der Kalibrierung auf ca.  $\pm 1$  cm nach der Kalibrierung reduziert werden. In (Lindner & Kolb 2006) werden die Distanzkorrekturterme ebenfalls über zuvor gemessene Referenzstrecken getrennt bestimmt. Die Distanzkalibrierung wird in eine auf die gesamte Sensorebene global wirkende Kalibrierung oder eine pixelweise Kalibrierung unterschieden. Beide Ansätze reduzieren zuvor aufgetretene Soll-Ist-Abweichungen auf unter einen Zentimeter. In (Lindner & Kolb 2007) kommt zusätzlich eine hochauflösende CCD-

Kamera (*Charge-Coupled Device*) im Zuge der Korrektur des amplitudenbasierten Distanzfehlers zum Einsatz. Fuchs und May (2007) sowie Fuchs und Hirzinger (2008) verwenden einen sehr präzise ansteuerbaren Roboterarm zur definierten Positionierung einer 3D-Kamera (Abbildung 10.3b). Im Gegensatz zu den bisher referierten Verfahren können signalwegeeffektbedingte und zyklische Distanzmessfehler nicht über den gesamten Messbereich von in der Regel ca. 7,5 m modelliert werden. Die Kalibrierung deckt hier lediglich 0,5 m bis 1,5 m ab. In Chiabrando u. a. (2009) wird eine ebene Zieltafel durch einen terrestrischen Laserscanner (TLS) koordinatenmäßig erfasst. Die zu kalibrierende 3D-Kamera wird anschließend auf dem TLS-Standpunkt zwangszentriert. Schiller u. a. (2008) verwenden fest mit der 3D-Kamera verbundene CCD-Kameras zur Herstellung des Bezugs zwischen dem Mehrsensorsystem und einem koordinatenmäßig bekannten Schachbrettmuster (Abbildung 10.3c).

Die Ermittlung von Referenzstrecken durch Messgeräte übergeordneter Genauigkeit ist in der Regel zeitlich und instrumentell sehr aufwendig. Diese Messungen umgehen Karel (2008), Robbins u. a. (2008; 2009) sowie Böhm und Pattinson (2010), indem sie den Datensatz (also die Amplituden- und Entfernungsbilder) und die Ergebnisse (also die Parameter der inneren und äußeren Orientierung) einer zuvor durchgeführten photogrammetrischen Referenzfeldkalibrierung verwenden. Somit können Sollstrecken zwischen den koordinatenmäßig bekannten Objektpunkten und den geschätzten Projektionszentren direkt berechnet und daraus die für das jeweils aufgestellte Fehlermodell notwendigen Korrekturterme abgeleitet werden.

### 10.3 Simultane photogrammetrische und Distanzkalibrierung

Bereits im vorangegangenen Abschnitt 10.2 wurde auf teils aufwendige Experimentalkonfigurationen für die Distanzkalibrierung hingewiesen und verschiedene Ansätze vorgestellt, welche die photogrammetrische und die Distanzkalibrierung zwar noch sequentiell, aber bereits unter der Verwendung des selben Datensatzes durchführen. Diese Ansätze können somit als quasi-simultan betrachtet werden, da die zuerst bestimmten Kameraorientierungen als Basis für die sich anschließende Distanzkalibrierung dienen.

Eine logische Erweiterung ist die simultane Berechnung aller Kalibrierwerte über die geeignete Integration der Beobachtungsarten in einem gemeinsamen Berechnungsschema. Dies ist aufgrund der Komplementarität der aus einem Messsignal rekonstruierten Amplituden- und Entfernungsinformationen vorteilhaft, um die Wechselwirkungen aller Modellparameter im funktionalen Zusammenhang zu berücksichtigen. Weiterhin kann bei einer geschlossenen Formulierung der gesamte stochastische Informationshaushalt des heterogenen Beobachtungsmaterials im Zuge einer automatischen Varianzkomponentenschätzung (VKS) ausgeschöpft werden. Auch bietet sich bspw. die Möglichkeit der Analyse von Parameterkorrelationen im Zuge der Fehlerrechnung an. Die Basis der geometrischen Modelle bilden zum einen die Kollinearität von im Amplituden- und/oder Entfernungsbild gemessenen Bildkoordinaten, ihren zugehörigen Objektpunkten und dem Projektionszentrum, zum anderen die euklidischen Distanzen zwischen dem Projektionszentrum und den Objektpunkten.

In (Westfeld 2007a) werden erste Überlegungen angestellt, die in Amplituden- und Entfernungsbildern gemessenen Bildkoordinaten geometrischer Primitive (Kugel, Pyramide) und die direkten Distanzmessungen in einem ausgleichenden Modell zu kombinieren (Bündelblockausgleichung mit festen Objektpunktkoordinaten). Lichti (2008) verwendet zwei in einem Abstand von 0,8 m installierte ebene, weiße Testfelder mit einer Größe von  $(3,6 \times 2,0) \text{ m}^2$ . Die Eckpunkte der angebrachten schwarzen, rechteckigen Zielmarken werden mit Kantenoperatoren in den Amplitudenbildern gemessen. Die zugehörigen Entfernungsmessungen müssen aufgrund von in der Regel nicht ganzzahligen Eckpunkten aus der nächsten Nachbarschaft interpoliert werden. Beide Beobachtungsarten werden zur Selbstkalibrierung der Aufnahmegeometrie einer SwissRanger<sup>TM</sup> SR3k und deren systematischen Distanzmessfehlern verwendet. Die verbleibenden Residuen der Entfernungsmessungen lagen nach der Ausgleichung im Mittel bei etwa 5 cm. Begründet werden kann dieses schlechte Ergebnis mit der ungünstig gewählten Experimentalkonfiguration. Die Interpolation der Entfernungswerte an den Eckpunkten schwarzer Zielmarken auf weißem Grund verursachen systematische Fehler durch Einflüsse von Reflektivität und *Scattering* (Abschnitte 9.7 und 9.3.2), welche Lichti nicht im Fehlermodell berücksichtigt. In (Lichti u. a. 2010) wird dieses Problem insoweit umgangen, dass nun ein einzelnes ca.  $(6,4 \times 4,0) \text{ m}^2$  großes, ebenes Testfeld mit weißen, kreisförmigen Zielmarken auf schwarzem Grund verwendet wird. Die Bildkoordinatenmessungen und die sich anschließenden Interpolationen der zugehörigen Entfernungswerte geschehen nun nicht mehr an den hochfrequenten Rändern, sondern in der homogenen Kreismitte (Abbildung 10.4a). Zwar werden, um den Mehrwegeeffekt zu verringern (Abschnitt 9.3), konvergente Aufnahmen aus unterschiedlichen Entfernungen erzeugt, aus diesen jedoch lediglich die Amplitudenbildkoordinatenmessungen in das Bündel integriert. Die zugehörigen Entfernungsmessungen stammen aus nur einer Orthogonalaufnahme, was eine Bestimmung des Maßstabsfaktors in Gleichung (9.1) ausschließt. Für die eingeführten Entfernungsmessungen erreichen Lichti u. a. (2010) eine mittlere quadratische Abweichung (*Root Mean Square Error, RMS*) von 8 mm.

Westfeld u. a. (2009) kalibrieren ebenfalls einen SwissRanger<sup>TM</sup> SR3k über die simultane Verwendung von im Amplituden- sowie Entfernungskanal gemessenen Bildkoordinaten und der direkt durch den Sensor ermittelten Entfernungsinformation. Bei diesem Ansatz wird ein Referenzfeld durch die sukzessive Verschiebung einer mit vier weißen Styroporkugeln signalisierten  $(60 \times 40) \text{ cm}^2$  Kalibrierplatte simuliert (Abbildung 10.4b). Der Durchmesser der Kugeln beträgt 10 cm. Im Amplitudenbild werden die Koordinaten mit 2D-LSM gemessen und als erste Beobachtungsart eingeführt (Abschnitt 7.6.1). Die auf den Distanzmessungen basierenden zwei weiteren Beobachtungsarten werden aus dem Entfernungsbild über eine auf dynamischer Schwellenwertbildung, RANSAC und Kugeleinpassung basierenden Kette von Bildanalysefunktionen extrahiert. Anhand der pro Kalibrierplattenstandpunkt erzeugten 3D-Kameraaufnahmen werden im Rahmen einer integrierten Bündelblockausgleichung mit festen Objektpunktkoordinaten die gesuchten Kalibrierwerte berechnet. Zur vollständigen Ausnutzung des Informationsgehalts werden die a-priori Varianzen der integrierten Beobachtungsarten während der Ausgleichung mitgeschätzt (Abschnitt 4.6). Die sehr große RMS-Abweichung der Entfernungsbeobachtungen von ca. 23 mm ist mit der Einführung von lediglich dem additiven

Distanzkorrekturterm  $d_0$  zu begründen (Gleichung 9.1).

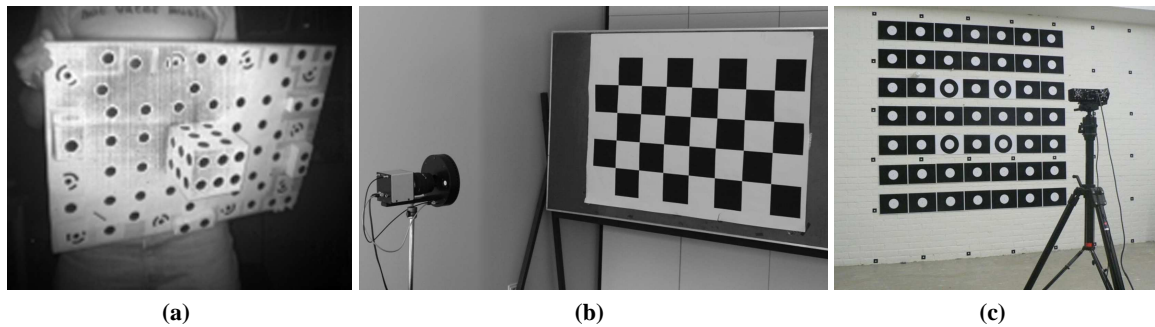
## 10.4 Diskussion

In den vorangegangenen Abschnitten wurden die bis jetzt auf 3D-Kameras adaptierten bzw. dafür neu entwickelten Kalibrierstrategien mit ihren Vor- und Nachteilen vergleichend vorgestellt. Sie unterscheiden sich vor allem hinsichtlich der verwendeten Kanäle (Amplituden-, Distanzbild), der Art und Weise wie die Beobachtungen zur Verfügung gestellt werden (direkte Messung oder Interpolation, durch die 3D-Kamera selbst oder einen zusätzlichen Sensor) und natürlich hinsichtlich der resultierenden Kalibrierparameter.

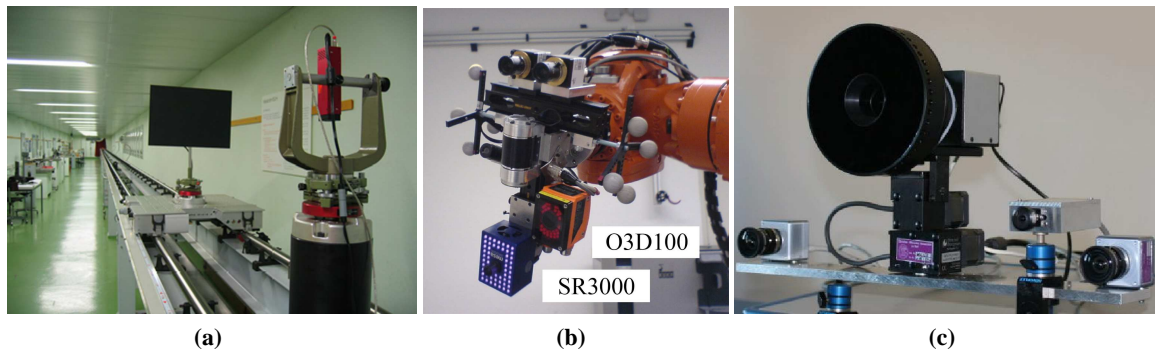
Als vom Autor besonders wichtig erachtet werden die Simultaneität und die Flexibilität eines 3D-Kamerakalibrierverfahrens sowie die Integration originärer Distanzbeobachtungen. Eine 3D-Kamera stellt aufgrund ihres inhärenten Designs und Messprinzips simultan Amplituden- und Entfernungsinformationen zur Verfügung, welche streng genommen aus einem einzigen Messsignal rekonstruiert werden (Abschnitt 8.2). Die Simultaneität ist somit die logische Forderung nach der Einbeziehung und der gemeinsamen Berücksichtigung aller verfügbaren Messkanäle einer 3D-Kamera im Rahmen eines integrierten Kalibrieransatzes. Im Gegensatz zu einer 2-Stufen-Kalibrierung können sich so die komplementären Eigenschaften der Beobachtungen gegenseitig stützen, ähnlich z. B. einer gemeinsamen Ausgleichung von Messwerten einer 2D-Kamera und eines Laserscanners (Schneider 2009). Die Verwendung zusätzlicher Sensorik (2D-Kamera, TLS) oder Messeinrichtungen (Roboterarm, Komparatorbank) zur Ermittlung einer Referenzgeometrie und die teilweise sehr große Dimensionierung des Referenzfelds erhöhen den experimentellen Aufwand enorm, erschweren die Reproduktion durch andere Arbeitsgruppen und stehen den Forderungen nach einem flexiblen photogrammetrischen 3D-Messverfahren entgegen. Die Bereitstellung von originären, also direkt gemessenen Beobachtungen bietet die Möglichkeit, die Einzelmessungen gut zu kontrollieren. Bei z. B. der Interpolation eines Entfernungswerts aus seiner direkten Nachbarschaft können vorhandene Ausreißer nicht erkannt werden. Auch ist dann die Angabe von a-priori und a-posteriori Genauigkeiten für eine einzelne ursprüngliche Distanzbeobachtung nicht möglich.

Das im folgenden Kapitel 11 vorgestellte Kalibrierverfahren soll möglichst alle genannten Eigenschaften in einem Ansatz vereinen.

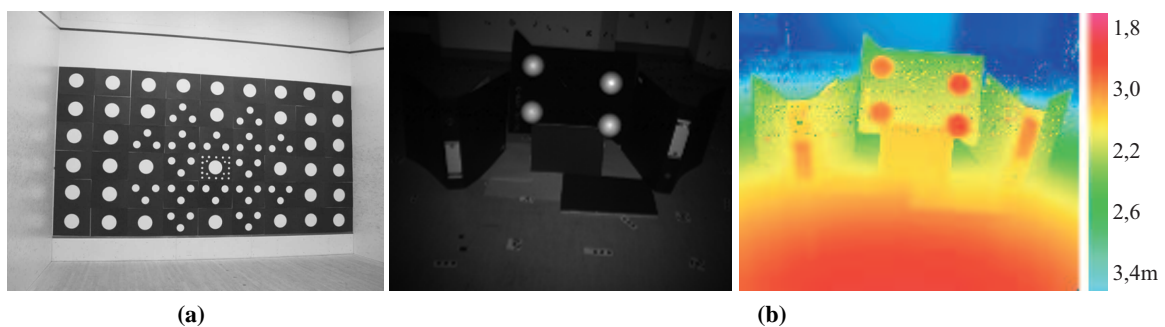




**Abb. 10.2:** Photogrammetrische 3D-Kamerakalibrierung: (a) Amplitudenbild eines im Rahmen der Arbeit erstellten 3D-Testfelds mit schwarzen kreisförmigen Zielmarken mit einem Durchmesser von 2 cm. (b) Experimentalkonfiguration mit einem in der *Computer Vision* üblichen ebenen Schachbrett-Testfeld (Beder & Koch 2008). (c) Spezielles ebenes Testfeld mit hochreflektierenden Zielmarken (Böhm & Pattinson 2010).



**Abb. 10.3:** Distanzkalibrierung einer 3D-Kamera: (a) SwissRanger™ SR-2 montiert auf einer automatischen interferometrischen Komparatorbank (Kahlmann 2007). (b) Industrieroboter KUKA KR 16 mit den 3D-Kameras SwissRanger™ SR3k und IFM O3D100 (Fuchs & Hirzinger 2008). (c) PMD-Kamera und zusätzliche CCD-Kameras zur gemeinsamen Kalibrierung (Schiller u. a. 2008).



**Abb. 10.4:** Simultane Kalibrierung einer 3D-Kamera: (a) Ebenes Testfeld an der Stirnwand eines Squashfelds (Lichti u. a. 2010). (b) Amplituden- und farbkodiertes Entfernungsbild einer 3D-Kalibrierplatte.



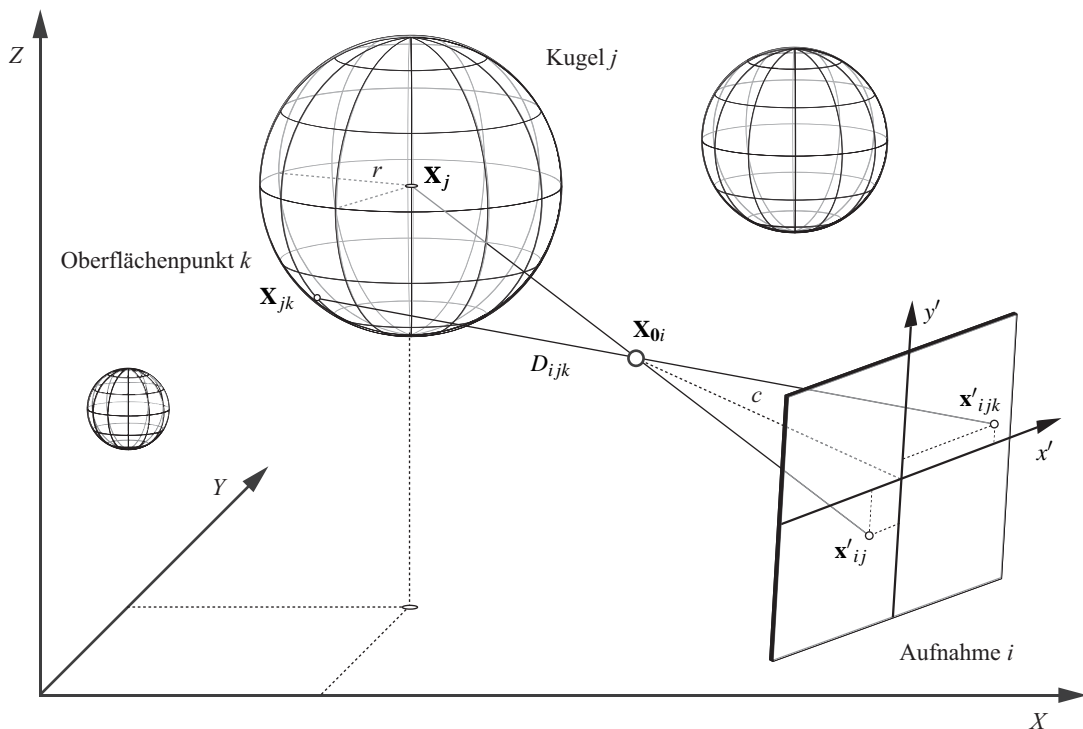
## Integrierte Bündelblockausgleichung

Die Abbildungsgeometrie einer 3D-Kamera entspricht einer Zentralprojektion (Abschnitt 5.1). Folglich müssen, wie bei einer konventionellen 2D-Kamera auch, Modellabweichungen zwischen dem Idealmodell der Zentralperspektive und der physikalischen Realität der Kamera über die Einführung entsprechender Korrekturparameter berücksichtigt werden (Abschnitte 5.1.3 und 10.1). Zusätzlich zur Amplitude erfassen 3D-Kameras Entfernungsinformationen für jedes Sensorelement. Die Ergebnisse der Entfernungsmessung werden ebenfalls durch diverse zufällige und systematische distanzmessspezifische Fehler beeinflusst (Kapitel 9). Ohne eine adäquate Modellierung kann es auch hier zu teils signifikanten Modellabweichungen kommen. Eine falsch berechnete Schrägdistanz wirkt sich bspw. auf die Reduktion auf eine Horizontalabstand und auf die Berechnung der kartesischen 3D-Objektcoordinate aus (Abschnitt 8.3).

Das folgende Kapitel stellt den aus (Westfeld 2007a; Westfeld u. a. 2009; Kapitel 10) weiterentwickelten integrierten Kalibrieransatz vor, welcher die folgenden Eigenschaften entsprechend den in Kapitel 2 gesetzten Zielen besitzen soll:

- Simultaneität: Das Hauptanliegen des Verfahrens stellt die Integration aller zur Verfügung stehenden 3D-Kamerainformationen in einem gemeinsamen funktionalen und stochastischen Modell dar.
- Authentizität: Die auszugleichenden Beobachtungen sind ursprünglichen Charakters, also nicht interpoliert, sondern direkt gemessen.
- Flexibilität: Zur Vermeidung teils aufwendiger Experimentalkonfigurationen basiert der zu entwickelnde Ansatz auf dem Prinzip einer Selbstkalibrierung. Die Verwendung zusätzlicher Sensoren höherer Genauigkeit zur Bestimmung von Referenzgeometrie oder eine Sensorfusion sind nicht notwendig.
- Adaptabilität: Die Parametrisierung des geometrischen Modells ist erweiterbar, z. B. um weitere auf die Distanzmessung wirkende Fehlereinflüsse. Die Integration geodätischer Zusatzinformationen zur Stabilisierung der Kalibrierung ist ebenfalls problemlos möglich.
- Konfiguration: Der Messaufbau erlaubt die Distanzkalibrierung über den gesamten Messbereich.
- Praktikabilität: Das Referenzfeld ist kompakt, portabel und einfach (de-)montierbar.

Bereits in der Einleitung des Kapitels 9 wurde darauf hingewiesen, dass das entwickelte integrierte Kalibrierverfahren eher als ein notwendiges Werkzeug denn als primärer Untersuchungsgegenstand dieser Arbeit zu verstehen ist. Somit musste im Folgenden auf einige Detailanalysen verzichtet werden, ein Verweis an entsprechender Stelle mit Vorschlägen für zukünftige Arbeiten wird gegeben.



**Abb. 11.1:** Geometrisches Modell der integrierten Bündelblockausgleichung: Für jede 3D-Kameraaufnahme  $i$  wird die Kugel  $j$  kreisförmig auf den Sensor projiziert. Für jeden Oberflächenpunkt  $k$  beinhaltet das Sensorelement  $(x', y')_{ijk}$  die Amplitudeninformation  $g_{v_{ijk}}$  und die Entfernungsmessung  $D_{ijk}$ .

## 11.1 Ausgleichungsansatz

Die primär umzusetzende Eigenschaft der integrierten Bündelblockausgleichung ist die simultane Ausgleichung von in beiden Informationskanälen beobachteten Größen zur gemeinsamen Schätzung aller Kalibrierwerte. Das Amplitudenbild soll, wie für ein konventionelles Intensitätsbild einer 2D-Kamera in der Photogrammetrie üblich, zur Messung von Bildkoordinaten von im Objektraum positionierten Referenzpunkten dienen. Das Entfernungsbild liefert zusätzlich für jedes diskrete Sensorelement Schrägdistanzen. Für die in der Regel nicht ganzzahlige im Amplitudenbild gemessene Position tut es das aber nicht. Die Herstellung des geometrischen Zusammenhangs zwischen diesen zwei Beobachtungsarten wäre somit nur über die Interpolation des Entfernungswerts möglich, was aber den zuvor definierten Eigenschaften des zu entwickelnden Kalibrierverfahrens widerspricht. Zur Vermeidung von Interpolationen im ursprünglichen Datenmaterial müssen geeignete Referenzpunkte gefunden werden, die sowohl eine direkte Integration der Entfernungswerte erlauben als auch einen geometrischen Zusammenhang zu den im Amplitudenbild gemessenen Bildkoordinaten liefern.

Als Referenzkörper (3D-Zielmarke) im Objektraum wird eine Kugel vorgeschlagen, welche als geometrisches Primitiv nur durch ihren Mittelpunkt und Radius analytisch beschrieben ist. Zudem sind Kugeln rotationsinvariant und werden aus jeder Sicht kreisförmig auf die Bildebene projiziert. Eine photogrammetrische Messung der Bildkoordinaten der Kugelmittelpunkte ist somit problemlos

möglich. Wie oben bereits beschrieben, liegt für den gemessenen Kugelmittelpunkt jedoch keine direkte Entfernungsmessung vor. Dennoch können die Distanzen über die Punkte der Kugeloberfläche in die Bündelblockausgleichung integriert werden. Das geometrische Modell basiert somit zum einen auf der Kollinearität des im Bild gemessenen Kugelmittelpunkts  $\mathbf{x}'_{ij}$ , dem Projektionszentrum  $\mathbf{X}_{0i}$  und dem korrespondierenden Objektpunkt  $\mathbf{X}_j$  (Abschnitt 5.1.5), zum anderen auf der euklidischen Distanz  $D_{ijk}$  zwischen dem Projektionszentrum und dem auf einem diskreten Sensorelement  $\mathbf{x}'_{ijk}$  abgebildeten Objektpunkt  $\mathbf{X}_{jk}$  der Kugeloberfläche (Abschnitt 8.3). Die Abbildung 11.1 verdeutlicht den aufgestellten Zusammenhang.

Für die Indizierung der eingeführten Variablen werden für die folgenden Kapitel die in Tabelle 11.1 gelisteten Vereinbarungen getroffen. Beispielsweise werden von einem koordinatenmäßig nicht bekannten Referenzfeld bestehend aus  $J$  Kugeln insgesamt  $I$  Aufnahmen erstellt. Aufgrund von z. B. Verdeckungen sind in jeder Aufnahme  $i$  nur  $J_i$  Kugeln sichtbar. Jede Kugel  $j$  wird durch ihre insgesamt  $K_{ij}$  abgebildeten Kugeloberflächenpunkte  $k$  repräsentiert. Die Summe aller im Bild  $i$  sichtbaren Kugeloberflächenpunkte beträgt  $K_i$ , für alle Bilder betrage sie  $K$ .

**Tab. 11.1:** Vereinbarungen zur Indizierung von Beobachtungen und Unbekannten.

Index	Erläuterung
$i$	Aufnahme (auch Standpunkt)
$I$	Anzahl aller Aufnahmen
$j$	Kugel (Mittelpunkt)
$J$	Anzahl aller Kugeln
$J_i$	Anzahl aller im Bild $i$ sichtbaren Kugeln
$k$	Kugeloberflächenpunkt
$K$	Anzahl aller Kugeloberflächenpunkte
$K_i$	Anzahl aller im Bild $i$ sichtbaren Kugeloberflächenpunkte aller Kugeln
$K_{ij}$	Anzahl aller im Bild $i$ sichtbaren Kugeloberflächenpunkte der Kugel $j$

Die Abbildung 11.2 stellt zusammenfassend den Datenfluss der entwickelten integrierten Bündelblockausgleichung dar. Die Eingabewerte sind die in den Amplitudenbildern gemessenen Bildkoordinaten der Kugelmittelpunkte und die durch die 3D-Kamera direkt gemessenen Distanzen zu einem beliebigen Kugeloberflächenpunkt (Abschnitt 11.2). Zusätzlich werden geometrische Bedingungen zur Erhöhung der Systemredundanz und zur Festlegung der Netzlagerung berücksichtigt (Abschnitt 11.4). Weiterhin müssen aufgrund der Nichtlinearität des funktionalen Modells Näherungswerte für alle unbekannt Parameter bereitgestellt werden (Abschnitt 11.3). Die für die Lösung der Ausgleichungsaufgabe in Abschnitt 11.7 notwendigen funktionalen und stochastischen Zusammenhänge werden in den zwei vorherigen Abschnitten 11.5 und 11.6 aufgezeigt. Für diese Arbeit sind vor allem die resultierenden Parameter der inneren Orientierung und die Distanzkorrekturterme der zu kalibrierenden 3D-Kamera von großem Interesse. Darüber hinaus sind die ausgeglichenen 3D-

Objektkoordinaten der Kugelmittelpunkte und der Kugeloberflächenpunkte sowie die Parameter der äußeren Orientierungen Ergebnisse der integrierten Bündelblockausgleichung. Letztlich liefert das Verfahren stochastische Informationen zur abschließenden Modellvalidierung und Qualitätsanalyse (Abschnitt 11.8).

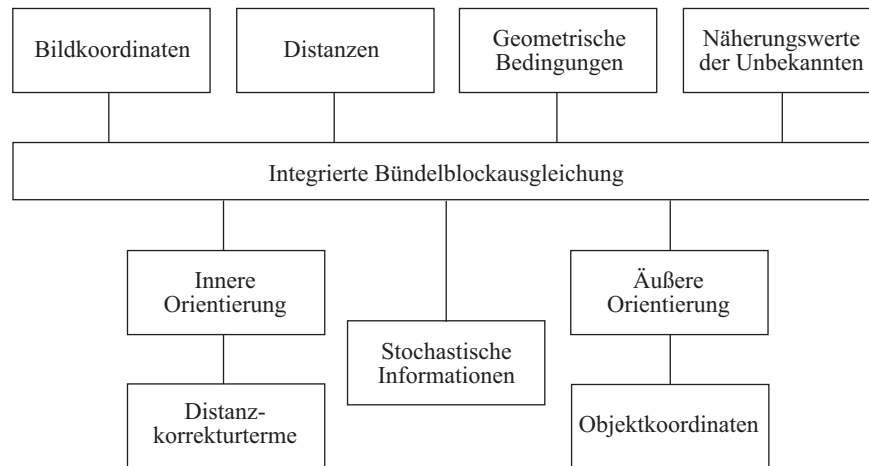


Abb. 11.2: Datenfluss für die integrierte Bündelblockausgleichung.

## 11.2 Bereitstellung der Beobachtungen

### 11.2.1 Beobachtungsgleichungen

Die im Amplitudenbild  $i$  für die sichtbare Kugel  $j$  automatisch gemessenen Bildkoordinaten  $\mathbf{x}'_{ij}$  der Mittelpunkte  $\mathbf{X}_j$  liefern jeweils zwei Beobachtungsgleichungen (Abschnitte 5.1.5 und 11.2.2):

$$\begin{aligned} \Phi_{ij}^{Ax'} : \quad x'_{ij} &= x'_0 - c \cdot \frac{r_{11i} \cdot (X_j - X_{0i}) + r_{21i} \cdot (Y_j - Y_{0i}) + r_{31i} \cdot (Z_j - Z_{0i})}{r_{13i} \cdot (X_j - X_{0i}) + r_{23i} \cdot (Y_j - Y_{0i}) + r_{33i} \cdot (Z_j - Z_{0i})} + \Delta x' \\ \Phi_{ij}^{Ay'} : \quad y'_{ij} &= y'_0 - c \cdot \frac{r_{12i} \cdot (X_j - X_{0i}) + r_{22i} \cdot (Y_j - Y_{0i}) + r_{32i} \cdot (Z_j - Z_{0i})}{r_{13i} \cdot (X_j - X_{0i}) + r_{23i} \cdot (Y_j - Y_{0i}) + r_{33i} \cdot (Z_j - Z_{0i})} + \Delta y' \end{aligned} \quad (11.1)$$

Als Funktion dieser Beobachtungen werden folgende in Abschnitt 11.3 detailliert eingeführte Unbekannte bestimmt:

- Innere Orientierung ( $c, \mathbf{x}'_0$ ) und in  $\Delta \mathbf{x}'$  enthaltene Zusatzparameter der verwendeten 3D-Kamera
- Äußere Orientierung ( $\mathbf{X}_{0i}, \omega_i, \varphi_i, \kappa_i$ ) für jede Aufnahme  $i$
- 3D-Objektkoordinaten  $\mathbf{X}_j$  für jeden Mittelpunkt der Kugel  $j$

Die direkt durch die 3D-Kamera beobachteten und automatisch aus den Entfernungsbildern segmentierten Schrägdistanzen  $D_{ijk}$  zwischen den Projektionszentren  $\mathbf{X}_{0i}$  und den Kugeloberflächen-

punkten  $\mathbf{X}_{jk}$  liefern jeweils eine weitere Beobachtungsgleichung (Abschnitte 8.3 und 11.2.3):

$$\Phi_{ijk}^D: D_{ijk} = \sqrt{(X_{jk} - X_{0i})^2 + (Y_{jk} - Y_{0i})^2 + (Z_{jk} - Z_{0i})^2} + \Delta D_{ijk} \quad (11.2)$$

Durch diese Beobachtungsgleichung werden die folgenden Unbekannten bestimmt:

- Distanzkorrekturterme  $d_{[0,6]}$
- Projektionszentrum  $\mathbf{X}_{0i}$  für jede Aufnahme  $i$
- 3D-Objektkoordinaten  $\mathbf{X}_{jk}$  für jeden Kugeloberflächenpunkt  $k$  der Kugel  $j$

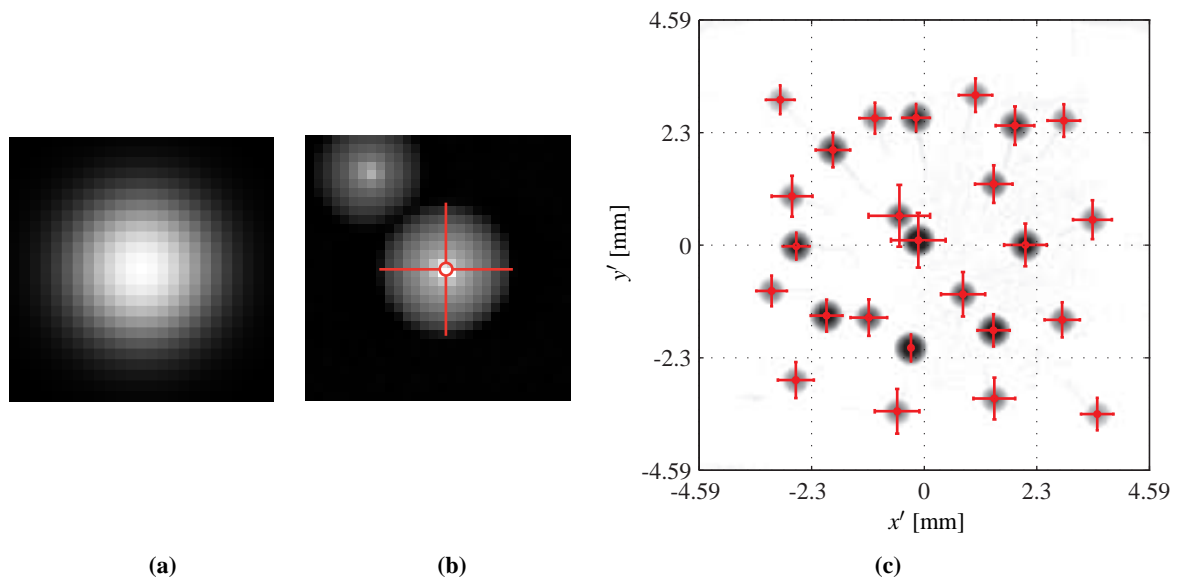
Mit dem bereits in Gleichung (11.1) eingeführten unbekanntem Projektionszentrum  $\mathbf{X}_0$  als ein Element der äußeren Orientierung ist in Kombination mit Gleichung (11.2) eine erste funktionale Verknüpfung der verschiedenen Beobachtungsarten umgesetzt. Sie wird durch die zusätzlich aufzustellenden Bedingungen zwischen den verbleibenden Unbekannten in Abschnitt 11.4 fortgesetzt.

### 11.2.2 Koordinatenmessung im Amplitudenbild

Die Mittelpunktkoordinaten  $\mathbf{x}'_{ij}$  der im Amplitudenbild  $i$  kreisförmig abgebildeten Kugeln  $j$  werden mit dem 2D-Korrelationsverfahren *Least Squares Template Matching* (LSM; Abschnitt 7.6) subpixelgenau gemessen und stellen die erste Beobachtungsart dar. Das hierfür notwendige generische Muster wird automatisch erzeugt und passt sich bereits vor der eigentlichen Bildzuordnung dynamisch an den Intensitätsverlauf und, unter Zuhilfenahme des zugehörigen Entfernungsbilds, an die Größe des Originals an (Abbildungen 11.3a und 11.3b).

Der iterative nichtlineare 2D-LSM Ansatz benötigt Näherungswerte für alle zu bestimmenden Affintransformationsparameter. Näherungswerte für die Translationsparameter  $(a_0, b_0)$  werden durch den von Burkhart (2007) implementierten Ansatz zur automatischen Messung kreisförmiger Zielmarken hinreichend genau bereitgestellt. Die Maßstabsparameter können aufgrund der oben genannten dynamischen *Template*-Erzeugung mit Eins initialisiert werden. Verbleibende radiometrische Unterschiede zwischen den Korrelationsfenstern werden vor der eigentlichen Anpassung nach der Methode der kleinsten Quadrate (MKQ) über eine lineare Histogrammanpassung korrigiert. Aufgrund einer kreisförmigen Kugelabbildung genügt ein reduzierter 2D-Affintransformationsparametersatz für die geometrische Transformation (Gleichung 7.18): Die Verschiebungen  $a_0$  in Zeilen- und  $b_0$  in Spaltenrichtung sowie die Verwendung eines einheitlichen Maßstabsparameters  $a_1 = b_2$  in beiden Koordinatenrichtungen sind ausreichend. Die Implementierung eines Signifikanztests ( $P = 95\%$ ; Abschnitt 4.5.6) verhindert darüber hinaus eine Überparametrisierung des Gleichungssystems für den Fall, dass die Abweichung zwischen dem Näherungswert und dem geschätzten Parameter das zufällige Messrauschen unterschreitet.

Exemplarisch für alle durchgeführten Messungen sind in Abbildung 11.3c die a-posteriori Standardabweichungen der Translationsparameter eines Bilds dargestellt. Im quadratischen Mittel betragen sie ca.  $1/25$  px, was bei ausreichender Überbestimmung ein typischer Wert für die erreichbare innere Genauigkeit ist (Baltsavias 1999). Die Abbildung 11.4 soll ferner darstellen, dass



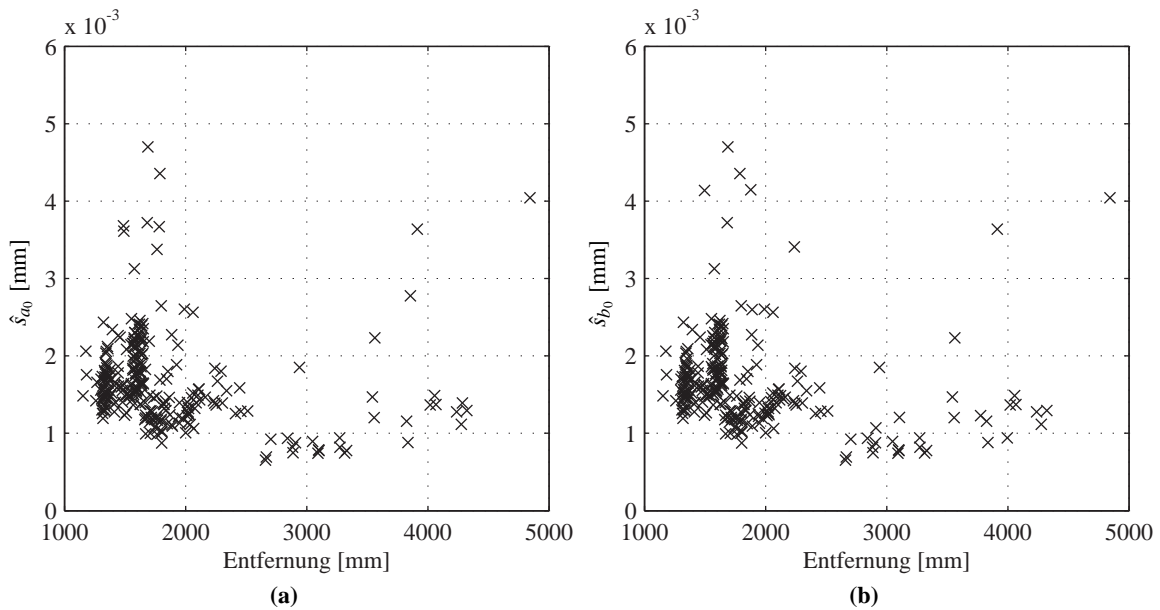
**Abb. 11.3:** Amplitudenbildpunktmesung mit 2D-LSM: (a) Automatisch erzeugtes Musterfenster. (b) Ausschnitt des 2D-affintransformierten Suchbilds mit Messkreuz. (c) Plot der a-posteriori Standardabweichungen  $\hat{s}$  der Translationsparameter  $(a_0, b_0)$  für ein invers hinterlegtes Amplitudenbild. Die mittlere quadratische Abweichung beträgt  $1,70 \mu\text{m}$  bei einer Pixelgröße von  $25 \mu\text{m}$ . Überhöhung: 200.

aufgrund der kreisförmigen Kugelabbildung erwartungsgemäß keine signifikanten Unterschiede zwischen den Koordinatenrichtungen auftreten. Auch kann keine Abhängigkeit der inneren LSM-Bildmessgenauigkeit von der Entfernung, also dem abgebildeten Kreisdurchmesser, erkannt werden.

Die gemessenen Bildkoordinaten werden als erste Beobachtungsart in das funktionale Modell der integrierten Bündelblockausgleichung eingeführt (Abschnitt 11.5). Die gelieferten Angaben zur Genauigkeit der Bildpunktmesung werden als Näherungswerte für die im Rahmen der Bündelblockausgleichung zu schätzenden a-priori Varianzkomponenten herangezogen (Abschnitt 11.6).

### 11.2.3 Kugelsegmentierung im Entfernungsbild

Die Entfernungsmessungen  $D_{ijk}(\mathbf{x}')$  zu jedem Oberflächenpunkt  $k$  einer Kugel  $j$  vom Standpunkt  $i$  stellen die zweite Beobachtungsart dar. Im Gegensatz zu den Bildkoordinatenmessungen in den Amplitudenbildern werden diese direkt vom Sensor für jedes Element gemessen (Abschnitt 8.2). Es besteht somit die Aufgabe, die beobachteten Entfernungswerte automatisch für jede Kugel in jedem Bild zu segmentieren, wobei zur Steigerung der Zuverlässigkeit neben den Entfernungsbildern auch die Amplitudenbilder Verwendung finden. Die Abbildung 11.5 zeigt das Ablaufschema, die Abbildung 11.6 visualisiert die Ergebnisse. Die einzelnen Segmentierungsstufen werden im Folgenden kurz besprochen.



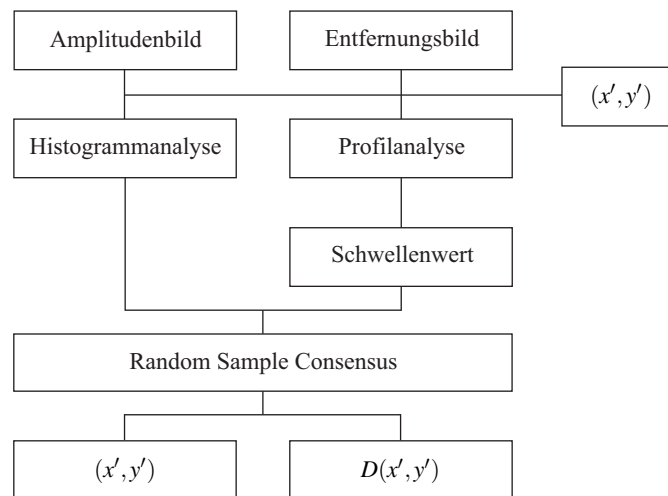
**Abb. 11.4:** A-posteriori Standardabweichungen  $\hat{s}_i$  der 2D-LSM Translationsparameter  $a_0$  und  $b_0$ . Aufgrund der kreisförmigen Kugelabbildung treten erwartungsgemäß keine Unterschiede zwischen den Koordinatenrichtungen auf. Auch kann keine Abhängigkeit der Bildmessgenauigkeit von der Entfernung, also dem abgebildeten Kreisdurchmesser, erkannt werden. RMS:  $\bar{s}_{a_0} = 1,70 \mu\text{m}$  und  $\bar{s}_{b_0} = 1,69 \text{e}^{-3} \text{mm}$  bei einer Pixelgröße von  $25 \mu\text{m}$ .

### Grobe Lokalisierung

Die Kugelmittelpunktkoordinaten  $\mathbf{x}'_{ij}$  werden durch 2D-LSM gemessen (Abschnitt 11.2.2). Die zugehörige Distanz  $D_{ij}$  wird im Entfernungsbild aufgrund einer in der Regel nicht ganzzahligen Bildkoordinatenmessung aus der nächsten Nachbarschaft interpoliert. Ausgehend von dem Kugelmittelpunkt findet eine erste grobe Ausmaskierung des Bildinhalts statt, wobei die Größe des Maskierungsfensters ( $dx', dy'$ ) mit dem im Objektraum bekannten Kugelradius  $R$  und der näherungsweise bekannten Kamerakonstante  $c$  über die Verhältnisgleichung  $dx' = dy' = 2 \cdot \frac{2R \cdot c}{D}$  gerechnet wird (Abbildungen 11.6a und 11.7).

### Histogrammanalyse im Amplitudenbild

Die in Abschnitt 11.8.1 vorgestellte Experimentalkonfiguration mit gut nahes Infrarotlicht (NIR) reflektierenden weißen Styroporkugeln erlaubt die Segmentierung des Amplitudenbilds mittels einer einfachen Histogrammanalyse. In Abbildung 11.8a ist das Histogramm für den in Abbildung 11.7 gezeigten Bildausschnitt dargestellt. Nach der Elimination aller Kandidaten mit einem Amplitudenwert kleiner als ein dynamischer Schwellenwert  $\varepsilon_A$  ist bereits eine Trennung zwischen Vorder- und Hintergrund zuverlässig möglich. Die Abbildungen 11.6b und 11.8b zeigen das Ergebnis.



**Abb. 11.5:** Schematische Darstellung der sequentiellen Kugelsegmentierung aus Amplituden- und Entfernungsbildern.

### Profilanalyse im Entfernungsbild

Das Entfernungsbild wird als zweiter wichtiger Segmentierungskanal verwendet. Wieder ausgehend vom durch **2D-LSM** bestimmten Kugelmittelpunkt erfolgt eine sternförmige Profilanalyse (**Abbildung 11.9**). Hierbei werden die Entfernungsunterschiede benachbarter Sensorpositionen entlang des Profilverlaufs berechnet. Überschreitet die Differenz einen zuvor festgelegten Schwellenwert vom Dreifachen des Kugelradius, kommt es also zu einem signifikanten Sprung in den Entfernungswerten, ist anzunehmen, dass der Rand der als Kreis abgebildeten Kugel erreicht ist. Für jedes Profil ergibt sich so ein Wert für die Kugelausdehnung in beide Profilrichtungen. Abschließend werden grobe Fehler, hervorgerufen durch z. B. weitere Kugeln oder durch die Kugelbefestigung, über eine Residuenanalyse eliminiert (**Abschnitt 4.5.4**). Alle Beobachtungen außerhalb des um den größten verbleibenden Radius verkleinerten Maskierungsfensters kommen als Kandidaten nicht mehr in Frage (**Abbildung 11.6c**).

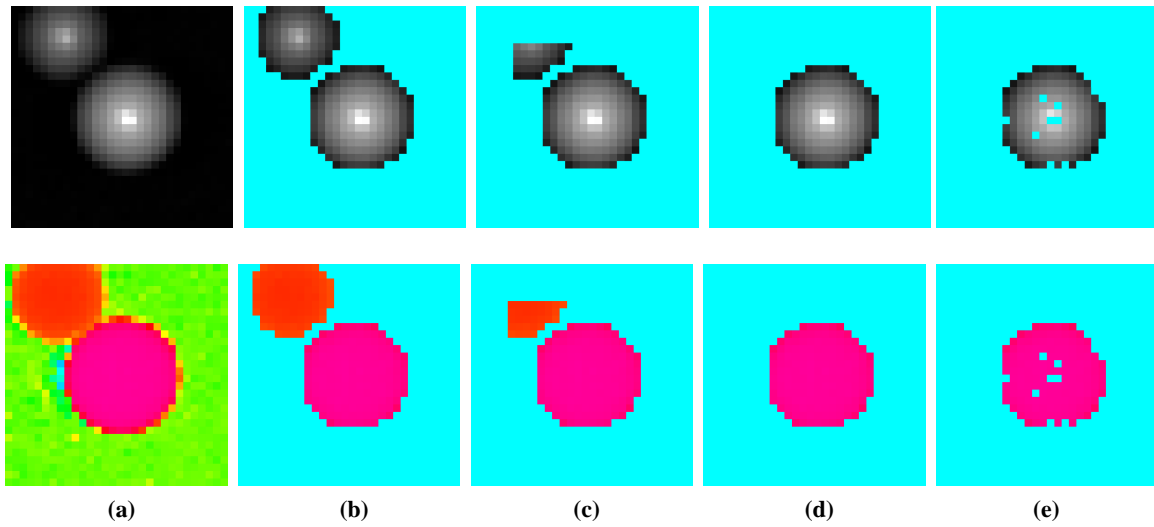
### Schwellenwertbildung im Entfernungsbild

Die direkt beobachtete Distanz des durch **2D-LSM** gemessenen Kugelmittelpunkts stellt die untere Schranke  $\varepsilon_{D1}$  eines dynamischen Schwellenwerts dar. Die obere Schranke  $\varepsilon_{D2}$  bildet der Entfernungswert des Rands der als Kreis abgebildeten Kugel, welche die oben beschriebene Profilanalyse zusätzlich liefert. Die dynamische Schwellenwertbildung eliminiert nun alle verbleibenden Kandidaten, die sich nicht im Entfernungsbereich der Kugelausdehnung befinden (**Abbildung 11.6d**).

### RANSAC

Die auf die Entfernungsmessungen einer **3D-Kamera** wirkenden Fehler beeinflussen selbstverständlich auch den oben diskutierten Segmentierungsprozess (**Kapitel 9**). Bevor die segmentierten Di-





**Abb. 11.6:** Visualisierung der Ergebnisse der einzelnen Programmstufen zur automatischen Kugelsegmentierung im Amplitudenbild (oben) und im farbkodierten Entfernungsbild (unten). Im Laufe der Segmentierung eliminierte Kandidaten werden Cyan dargestellt. (a) Grobe Lokalisierung. (b) Histogrammanalyse im Amplitudenbild. (c) Profilanalyse im Entfernungsbild. (d) Schwellenwertbildung im Entfernungsbild. (e) RANSAC.

stanzmessungen in die Bündelblockausgleichung integriert werden, müssen verbleibende Effekte, hervorgerufen durch z. B. Mehrwegeeffekte am Kugelrand oder Sättigungseffekte aufgrund direkter Signalreflexionen, abgefangen werden. Die Anzahl grober Fehler im Beobachtungsmaterial wird dadurch weiter minimiert und folglich die Ausgleichung stabilisiert.

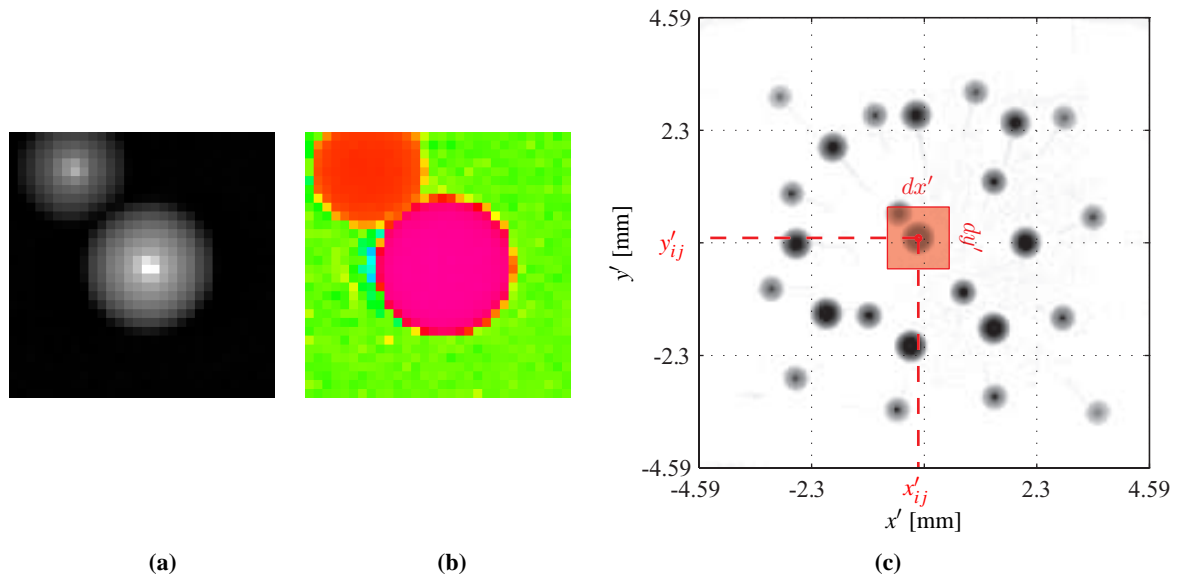
Hier bietet sich der Einsatz des robusten Schätzers **RANSAC** (*Random Sample Consensus*; Abschnitt 4.7) an. Per Zufallsgenerator werden vier Kugelpunkte ausgewählt und zur Berechnung einer geschlossenen Lösung des Kugelmodells verwendet:

$$0 = (x'_{ijk} - x'_{ij})^2 + (y'_{ijk} - y'_{ij})^2 + (d'_{\perp ijk} - d'_{\perp ij})^2 - (r'_{ij})^2 \quad (11.3)$$

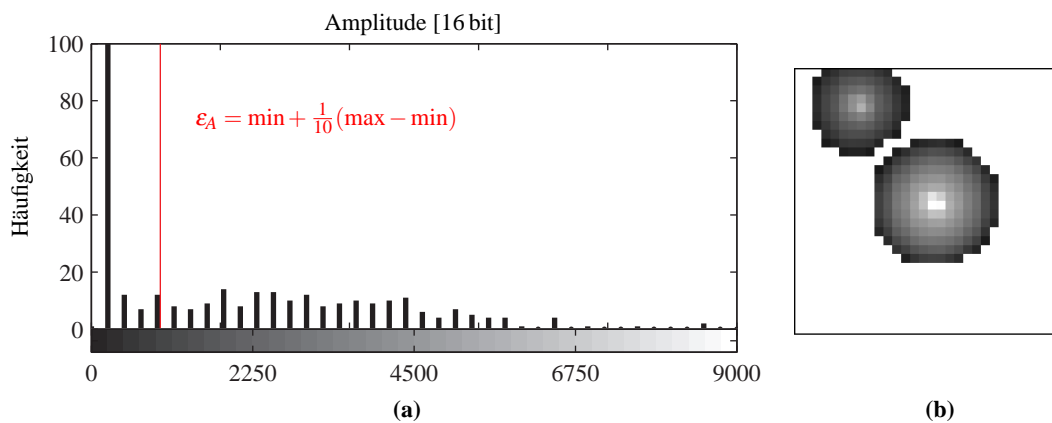
mit

$\mathbf{x}'_{ij}, d'_{\perp ij}$	Bildkoordinaten und auf das Sensorformat skalierte Horizontaldistanz des Mittelpunkts der Kugel $j$ im Bild $i$
$r'_{ij}$	Radius der Kugel $j$ im Bild $i$
$\mathbf{x}'_{ijk}$	Diskrete Sensorposition des Oberflächenpunkts $k$ auf der Kugel $j$ im Bild $i$
$d'_{\perp ijk}$	Zu $\mathbf{x}'_{ijk}$ zugehörige und auf das Sensorformat skalierte Horizontaldistanz

Anschließend werden der Teilmenge  $S_i$  diejenigen verbleibenden Beobachtungen hinzugefügt, welche das Kugelmodell (11.3) innerhalb einer empirisch ermittelten Fehlerschranke  $\varepsilon$  unterstützen. Der Algorithmus akzeptiert nach 72 Iterationen ( $z = 0,99$ ;  $w = 0,5$ ;  $n = 4$ ; vgl. Abschnitt 4.7) die Teilmenge mit der höchsten Anzahl unterstützender Beobachtungen, mindestens jedoch  $N_S = 5$  (Abbildung 11.6e).



**Abb. 11.7:** Basierend auf den 2D-LSM Bildkoordinaten  $(x', y')_{ij}$  der Kugelmittelpunkte werden im Amplituden- und Entfernungsbild (a,b) nicht relevante Bildinhalte grob ausmaskiert (c).



**Abb. 11.8:** Die Festlegung eines dynamischen Amplitudenschwellenwerts  $\epsilon_A$  erfolgt mittels einer einfachen Histogrammanalyse: (a) Histogramm. (b) Segmentiertes Amplitudenbild.

Als Alternative zu dem vorgestellten RANSAC-Algorithmus können die mit 2D-LSM gemessenen Bildkoordinaten der Kugelmittelpunkte und die aus der Profilanalyse resultierenden Bildradien direkt zur Lösung des Kugelmodells verwendet und anschließend die Differenzen zwischen den Beobachtungen und den Funktionswerten analysiert werden (Abschnitt 4.5.4).

#### 11.2.4 Anzahl der Beobachtungen

Die Gesamtanzahl  $n$  der Beobachtungen setzt sich zusammen aus der Anzahl  $n_A$  aller in den Amplitudenbildern gemessenen Bildkoordinatenpaare  $(x', y')_{ij}$  und aus der Anzahl  $n_D$  aller direkt gemessene-

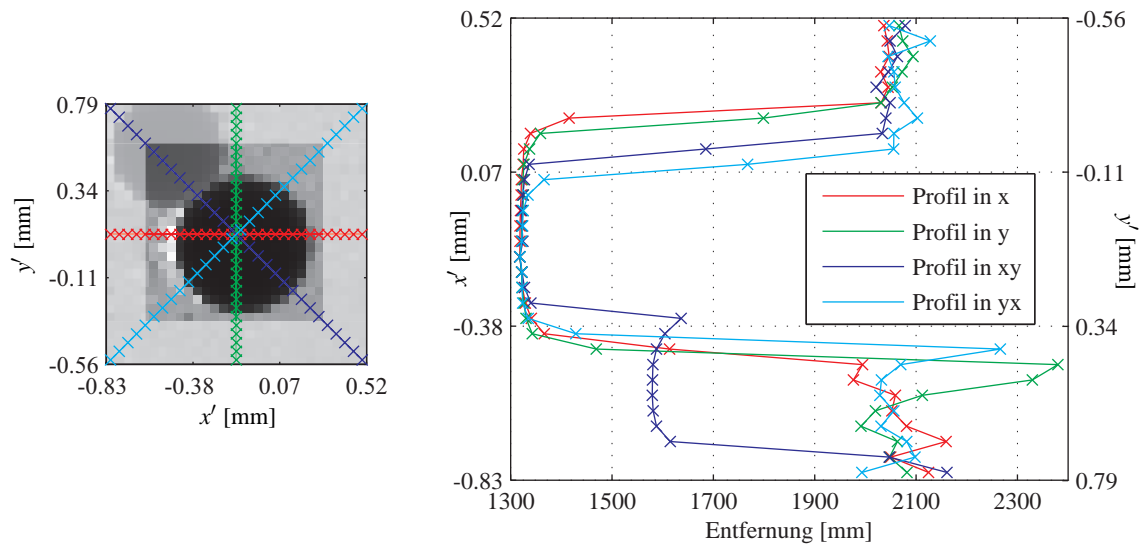


Abb. 11.9: Sternförmige Profilanalyse im Entfernungsbild.

nen Distanzen  $D_{ijk}$ :

$$n = n_A + n_D = \sum_{i=1}^I \left( 2 \cdot J_i + \sum_{j=1}^{J_i} K_{ij} \right) \quad (11.4)$$

Das folgende in [Tabelle 11.2](#) aufgeführte synthetische Beispiel soll für insgesamt  $I = 15$  Standpunkte den Beobachtungshaushalt verdeutlichen. Zur Wahrung der Übersichtlichkeit wird angenommen, dass von jedem Standpunkt  $i$  alle  $J = J_i = 25$  Kugeln sichtbar sind und, dass jede Kugel  $j$  durch genau  $K = K_{ij} = 30$  gemessene Kugeloberflächenpunkte repräsentiert wird.

Tab. 11.2: Synthetisches Beispiel zur Ermittlung der Anzahl der Beobachtungen.

Kanal	$I$	$J_i$	$K_{ij}$	$n$
Amplitudenbild	15	25	$\sum^I 2 \cdot J_i$	750
Entfernungsbild	15	25	30 $\sum^I \sum^{J_i} K_{ij}$	11.250
Gesamtanzahl der Beobachtungen				12.000

## 11.3 Erläuterung der Unbekannten und Beschaffung von Näherungswerten

### 11.3.1 Innere Orientierung und Zusatzparameter

Die Parameter der inneren Orientierung beschreiben den Zusammenhang zwischen dem kamerafesten Bildkoordinatensystem und dem übergeordneten Objektkoordinatensystem sowie mögliche Abweichungen vom Idealmodell der Zentralperspektive. Entsprechend dem [Abschnitt 5.1.3](#) setzen sich die Parameter der inneren Orientierung und von Bildfehler beschreibenden Funktionen wie folgt zusammen:

- Kamerakonstante  $c$
- Bildhauptpunkt  $\mathbf{x}'_0(x'_0, y'_0)$
- Parameter von Bildfehler beschreibenden Funktionen  $A_1, A_2, A_3, B_1, B_2, C_1, C_2$

Für die im Rahmen der Arbeit kalibrierten 3D-Kameras wird angenommen, dass sich ihre Abbildungsgeometrie während der Standpunktwechsel nicht signifikant ändert. Somit werden pro Experimentalkonfiguration maximal  $u_{IO} = 10$  bildinvariante Parameter der inneren Orientierung und Zusatzparameter im Rahmen der integrierten Bündelblockausgleichung geschätzt. Gehen einzelne Parameter als Konstante ein, verringert sich die Anzahl entsprechend.

Der Näherungswert der Kamerakonstante wird über die Herstellerangabe bereitgestellt ([MESA 2010](#); [PMDTec 2010](#)). Für die Position des Bildhauptpunkts wird zunächst angenommen, dass er mit der Sensormitte zusammenfällt. Der Bildfehlereinfluss wird bei der Linearisierung der Beobachtungsgleichungen für die erste Iteration vernachlässigt.

### 11.3.2 Äußere Orientierung

Die folgenden sechs Parameter der äußeren Orientierung legen die räumliche Lage und die Ausrichtung einer 3D-Kameraaufnahme  $i$  im übergeordneten Objektkoordinatensystem fest ([Abschnitt 5.1.4](#)):

- Projektionszentrum  $\mathbf{X}_{0i}(X_0, Y_0, Z_0)$
- Rotationswinkel  $\omega_i, \varphi_i, \kappa_i$

Bei insgesamt  $I$  Standpunkten beträgt die Anzahl  $u_{EO}$  der äußeren Orientierungselemente  $6 \cdot I$  und erhöht sich auf  $7 \cdot I$ , sobald, wie in dieser Arbeit geschehen, die Eulerschen Rotationswinkel  $(\omega, \varphi, \kappa)_i$  zur Vermeidung von Mehrdeutigkeiten durch die Quaternionen  $(q_1, q_2, q_3, q_4)_i$  ersetzt werden ([Abschnitt 4.3](#)).

Mit dem DLT-Ansatz ([Direkte lineare Transformation](#); [Abschnitt 5.3.2](#)) werden die Orientierungselemente aller Aufnahmen in einem linearen Gleichungssystem berechnet und als Näherungen für die folgende integrierte Bündelblockausgleichung verwendet.

### 11.3.3 Distanzkorrektur

Die Distanzmessung einer 3D-Kamera unterliegt verschiedenen Fehlereinflüssen (Kapitel 9). Handelt es sich um systematisch wirkende Fehler (Abschnitt 4.5.1) und ist der Fehlereinfluss modellierbar, kann das geometrische Modell in geeigneter Weise erweitert werden. Entsprechend den einleitenden Ausführungen im Kapitels 9 werden nur die distanzmessspezifischen Fehler in der integrierten Bündelblockausgleichung berücksichtigt, deren Ursachen physikalisch begründet und streng modellierbar sind. Hierzu gehören:

- Lineare Distanzkorrekturparameter  $d_0, d_1$
- Zyklische Distanzkorrekturparameter  $d_2, d_3, d_4, d_5$
- Signalwegeffektbedingter Distanzkorrekturparameter  $d_6$

Der Einfluss der Reflektivität und deren Korrektur sind zwar ebenfalls für eine Integration ausreichend beschrieben (Abschnitt 9.7), jedoch aufgrund der gewählten Experimentalkonfiguration mit ausschließlich weißen Styroporkugeln nicht signifikant bestimmbar (Abschnitt 11.8.1). Des Weiteren ist die entwickelte integrierte Bündelblockausgleichung insofern flexibel, als dass das funktionale Modell problemlos erweiterbar ist, sobald neue Vorschriften zur Modellierung von auf die Distanzmessung einer 3D-Kamera wirkenden Fehler vorliegen.

Aufgrund eines für 3D-Kameras signifikanten distanzbeeinflussenden linearen Trends (Abschnitt 9.5.1) müssen für die Distanzkorrekturwerte  $d_0$  und  $d_1$  hinreichend genaue Näherungswerte bereitgestellt werden. Deren Bestimmung ist über ein einfaches polares Anhängen der beobachteten Distanzen an das entsprechende Projektionszentrum unter Kenntnis der Näherungswerte der inneren und äußeren Orientierungsparameter prinzipiell möglich (Abschnitte 11.3.1), jedoch liegen zu diesem Zeitpunkt die hierfür notwendigen 3D-Objektkoordinaten der Kugeloberflächenpunkte noch nicht vor. Alternativ werden sie in dieser Arbeit ausgleichend über den Schnitt des näherungsweise rekonstruierten Projektionsstrahls mit der Kugeloberfläche geschätzt (Abschnitt 4.4.2). Durch das Einsetzen der 3D-Helmerttransformationsgleichung (8.6)

$$\mathbf{X}_{jk} = \mathbf{X}_{0i} + (D + \Delta D) \cdot \mathbf{R}_i \cdot \boldsymbol{\xi}' \quad (11.5)$$

in eine Kugelgleichung der allgemeinen Form

$$0 = (X_{jk} - X_j)^2 + (Y_{jk} - Y_j)^2 + (Z_{jk} - Z_j)^2 - r^2 \quad (11.6)$$

ergibt sich die hierfür notwendige Beobachtungsgleichung:

$$\begin{aligned} 0 = & [X_{0i} + (D_{ijk} + \Delta D_{ijk}) \cdot (r_{11}\xi'_1 + r_{12}\xi'_2 - r_{13}\xi'_3) - X_j]^2 \\ & + [Y_{0i} + (D_{ijk} + \Delta D_{ijk}) \cdot (r_{21}\xi'_1 + r_{22}\xi'_2 - r_{23}\xi'_3) - Y_j]^2 \\ & + [Z_{0i} + (D_{ijk} + \Delta D_{ijk}) \cdot (r_{31}\xi'_1 + r_{32}\xi'_2 - r_{33}\xi'_3) - Z_j]^2 - r^2 \end{aligned} \quad (11.7)$$

mit

$$\xi' = \|\mathbf{x}'\| = \left\| \begin{bmatrix} x' - x'_0 - \Delta x' \\ y' - y'_0 - \Delta y' \\ -c \end{bmatrix} \right\|$$

Alle weiteren Distanzkorrekturterme  $d_2$  bis  $d_6$  werden mit Null initialisiert.

### 11.3.4 3D-Objektpunktkoordinaten

Bei der vorgestellten Methode handelt es sich um eine Bündelblockausgleichung mit Selbstkalibrierung (Abschnitt 5.4.3). Im Gegensatz zu einer Referenzfeldkalibrierung mit einem koordinatenmäßig bekannten Testfeld (Abschnitt 5.3.3) werden die folgenden zwei Arten von 3D-Objektkoordinaten mitgeschätzt:

- 3D-Objektkoordinaten der Kugelmittelpunkte  $\mathbf{X}_j(X, Y, Z)$
- 3D-Objektkoordinaten der Kugeloberflächenpunkte  $\mathbf{X}_{jk}(X, Y, Z)$

Die 3D-Objektkoordinaten der Kugelmittelpunkte resultieren aus Strahlenschnitten von in Amplitudenbildern gemessenen Bildkoordinaten der kreisförmig abgebildeten Kugeln (Abschnitt 11.2.2); eine direkte Distanzmessung liegt nicht vor. Die 3D-Objektkoordinaten der Kugeloberflächenpunkte werden im Gegensatz dazu einzig über polares Anhängen der beobachteten Distanzwerte berechnet (Abschnitte 8.3 und 11.2.3). Eine Rekonstruktion über Strahlenschnitte ist nicht möglich, da für die diskreten und ganzzahligen Sensorpositionen in der Regel keine homologen Punkte in den anderen Bildern existieren.

Die Näherungskordinaten der Kugelmittelpunkte entstammen dem Herstellungsprozess der Kalibrierplatte (Abschnitt 11.8). Mit den in Abschnitt 11.3.3 näherungsweise bereitgestellten Distanzkorrekturtermen können die Näherungswerte der 3D-Objektkoordinaten der Kugeloberflächenpunkte über polares Anhängen mittels Gleichung (8.6) berechnet werden.

### 11.3.5 Anzahl der unbekannt Parameter

Die maximale Anzahl  $u$  der unbekannt Parameter ist die Summe der Anzahl der inneren Orientierungselemente  $u_{IO}$ , der äußeren Orientierungselemente  $u_{EO}$ , der Distanzkalibrierparameter  $u_D$  sowie der Objektpunktanzahl  $u_{\mathbf{X}_j}$  und  $u_{\mathbf{X}_{jk}}$ :

$$\begin{aligned} u &= u_{IO} + u_{EO} + u_D + u_{\mathbf{X}_j} + u_{\mathbf{X}_{jk}} \\ &= 10 + 7 \cdot I + 7 + \sum_{i=1}^I \left( 3 \cdot J_i + \sum_{j=1}^{J_i} 3 \cdot K_{ij} \right) \end{aligned} \quad (11.8)$$

Zur Verdeutlichung soll das für die Beobachtungsanzahl in Tabelle 11.2 begonnene synthetische Beispiel für die maximale Anzahl der unbekannt Parameter in Tabelle 11.3 fortgesetzt werden.

**Tab. 11.3:** Synthetisches Beispiel zur Ermittlung der maximalen Anzahl an unbekanntem Parametern.

Unbekannte Parameter			$I$	$J_i$	$K_{ij}$	$u$
Innere Orientierung	$c$	Kamerakonstante				1
	$\mathbf{x}'_0$	Bildhauptpunkt				2
	$\Delta \mathbf{x}'$	Abbildungsfehler				7
Äußere Orientierung	$\mathbf{X}_0$	Projektionszentrum	15		$3 \cdot I$	45
	$q_1, q_2, q_3, q_4$	Rotationselemente	15		$4 \cdot I$	60
Distanzkalibrierung	$d_0, d_1$	Linearer Anteil				2
	$d_2, d_3, d_4, d_5$	Zyklischer Anteil				4
	$d_6$	Latenzbasierter Anteil				1
3D-Objektpunkt	$\mathbf{X}$	Kugelmittelpunkt	15	25	$\sum^I 3 \cdot J_i$	1.125
		Kugeloberflächenpunkt	15	25	30	$\sum^I \sum^{J_i} 3 \cdot K_{ij}$
Gesamtanzahl der unbekanntem Parameter						34.997

## 11.4 Zusätzliche Bedingungen

Die Einführung einer zusätzlichen Beobachtungsgleichung (11.2) pro Entfernungsmessung schließt auch die Bestimmung der drei unbekanntem Objektpunktkoordinaten  $(X, Y, Z)_{jk}$  der Kugeloberfläche mit ein. Einer neuen Beobachtungsgleichung stehen folglich drei neue Unbekannte gegenüber. Für das in den letzten beiden Abschnitten 11.2 und 11.3 begonnene synthetische Beispiel führt dies nach Gleichung (4.36) zu einer für das Gleichungssystem negativen Anzahl an Freiheitsgraden. Das so entstandene unbestimmte Gleichungssystem kann nur durch die Integration weiterer Gleichungen in ein überbestimmtes Gleichungssystem zurückgeführt und anschließend ausgleichend gelöst werden. Entsprechend den Ausführungen der Abschnitte 4.4.3 und 5.4.4 können bspw. zusätzliche Bedingungsgleichungen zwischen den Unbekanntem in Form mathematisch-geometrischer Relationen formuliert werden. Dies erhöht einerseits die Systemredundanz, andererseits wird die in Abschnitt 11.2 begonnene direkte und indirekte funktionale Verknüpfung der 3D-Kameraentfernungsmessungen mit möglichst allen unbekanntem Parametern gewährleistet.

### 11.4.1 Kugelbedingung

Als eine erste, im Ausgleichungsprozess zu erfüllende, Restriktion bietet sich die Berücksichtigung der kugelförmigen Geometrie der Referenzkörper an. Alle Oberflächenpunkte  $\mathbf{X}_{jk}$  einer Kugel  $j$  mit ihrem Mittelpunkt  $\mathbf{X}_j$  sowie ihrem konstanten und a-priori bekannten Radius  $R$  haben die folgende Kugelgleichung zu erfüllen, welche insgesamt  $b_{Kugel} = n_D$  mal aufgestellt werden kann:

$$\Gamma_{jk}^{Kugel} : 0 = (X_{jk} - X_j)^2 + (Y_{jk} - Y_j)^2 + (Z_{jk} - Z_j)^2 - R^2 \quad (11.9)$$

### 11.4.2 Kollinearitätsbedingung

Weiterhin können pro Kugeloberflächenpunkt zwei Gleichungen, insgesamt  $b_{Koll}$ , integriert werden, die die Kollinearität von einem diskreten Bildkoordinatenpaar  $\mathbf{x}'_{ijk}$ , dem Projektionszentrum  $\mathbf{X}_{0i}$  und dem korrespondierenden Objektpunkt  $\mathbf{X}_{jk}$  der Kugeloberfläche fordern ([Abschnitt 5.1.5](#); [Abbildung 11.1](#)):

$$\begin{aligned} \Gamma_{ijk}^{Kollx'} : \quad 0 &= x'_0 - c \cdot \frac{r_{11i} \cdot (X_{jk} - X_{0i}) + r_{21i} \cdot (Y_{jk} - Y_{0i}) + r_{31i} \cdot (Z_{jk} - Z_{0i})}{r_{13i} \cdot (X_{jk} - X_{0i}) + r_{23i} \cdot (Y_{jk} - Y_{0i}) + r_{33i} \cdot (Z_{jk} - Z_{0i})} + \Delta x' - x'_{ijk} \\ \Gamma_{ijk}^{Kolly'} : \quad 0 &= y'_0 - c \cdot \frac{r_{12i} \cdot (X_{jk} - X_{0i}) + r_{22i} \cdot (Y_{jk} - Y_{0i}) + r_{32i} \cdot (Z_{jk} - Z_{0i})}{r_{13i} \cdot (X_{jk} - X_{0i}) + r_{23i} \cdot (Y_{jk} - Y_{0i}) + r_{33i} \cdot (Z_{jk} - Z_{0i})} + \Delta y' - y'_{ijk} \end{aligned} \quad (11.10)$$

### 11.4.3 Orthonormalitätsbedingung

Die aus numerischer Sicht sinnvolle Substitution der Eulerschen Rotationswinkel  $(\omega, \varphi, \kappa)_i$  durch vier Quaternionen  $(q_1, q_2, q_3, q_4)_i$  bedingt die Einführung einer zusätzlichen Gleichung pro Kamerastandpunkt  $i$ , welche die Orthonormalität der Rotationsmatrix verlangt ([Abschnitte 4.3](#) und [11.3.2](#)). Insgesamt werden  $b_{Quat} = I$  Gleichungen der folgenden Form aufgestellt:

$$\Gamma_i^{Quat} : \quad 0 = q_{1i}^2 + q_{2i}^2 + q_{3i}^2 + q_{4i}^2 - 1 \quad (11.11)$$

### 11.4.4 Datumsfestlegung

Letztlich erfordert die Datumsfestlegung die Einführung weiterer Nebenbedingungen. Entsprechend den Ausführungen des [Abschnitts 5.4.2](#) erfolgt die Lösung des Datumsproblems durch eine freie Netzlagerung mit drei Translationen

$$\begin{aligned} \Gamma_1^{Dat} : \quad 0 &= \sum^I dX_{0i} + \sum^J \left( dX_j + \sum^K dX_{jk} \right) \\ \Gamma_2^{Dat} : \quad 0 &= \sum^I dY_{0i} + \sum^J \left( dY_j + \sum^K dY_{jk} \right) \\ \Gamma_3^{Dat} : \quad 0 &= \sum^I dZ_{0i} + \sum^J \left( dZ_j + \sum^K dZ_{jk} \right), \end{aligned} \quad (11.12)$$



drei Rotationen

$$\begin{aligned}
\Gamma_4^{Dat} : \quad 0 &= \sum^I (-Z_{0i}dY_{0i} + Y_{0i}dZ_{0i}) + \sum^J \left( -Z_jdY_j + Y_jdZ_j + \sum^K (-Z_{jk}dY_{jk} + Y_{jk}dZ_{jk}) \right) \\
\Gamma_5^{Dat} : \quad 0 &= \sum^I (Z_{0i}dX_{0i} - X_{0i}dZ_{0i}) + \sum^J \left( Z_jdX_j - X_jdZ_j + \sum^K (Z_{jk}dX_{jk} - X_{jk}dZ_{jk}) \right) \\
\Gamma_6^{Dat} : \quad 0 &= \sum^I (-Y_{0i}dX_{0i} + X_{0i}dY_{0i}) + \sum^J \left( -Y_jdX_j + X_jdY_j + \sum^K (-Y_{jk}dX_{jk} + X_{jk}dY_{jk}) \right)
\end{aligned} \tag{11.13}$$

und einer Maßstabsanpassung

$$\begin{aligned}
\Gamma_7^{Dat} : \quad 0 &= \sum^I (X_{0i}dX_{0i} + Y_{0i}dY_{0i} + Z_{0i}dZ_{0i}) \\
&+ \sum^J \left( X_jdX_j + Y_jdY_j + Z_jdZ_j + \sum^K (X_{jk}dX_{jk} + Y_{jk}dY_{jk} + Z_{jk}dZ_{jk}) \right)
\end{aligned} \tag{11.14}$$

Diese sehr flexible Vorgehensweise ohne die Kenntnis von Passpunktinformationen ist möglich, da für die Bestimmung der 3D-Kamerakorrekturwerte die Rekonstruktion des Referenzfelds in einem beliebig gelagerten Objektkoordinatensystem erfolgen kann.

#### 11.4.5 Anzahl der zusätzlichen Bedingungen

Alle für jeden Punkt der Kugeloberfläche aufgestellten Nebenbedingungen (11.9) und (11.10) führen keine neuen unbekannt Parameter ein, sondern erhöhen in vollem Umfang den Redundanzanteil des damit wieder überbestimmten Gleichungssystems um  $b_{Kugel}$  und  $b_{Koll}$ . Weiterhin steigern die Orthogonalitätsbedingungen der Rotationsmatrizen die Anzahl der Bedingungsgleichungen um  $b_{Quat}$  und die Datumsfestlegung um  $b_{Dat}$ . Die Gesamtanzahl  $b$  lautet somit:

$$\begin{aligned}
b &= b_{Kugel} + b_{Koll} + b_{Quat} + b_{Dat} \\
&= n_D + 2 \cdot n_D + I + 7
\end{aligned} \tag{11.15}$$

Das synthetische Beispiel zur Verdeutlichung des Haushalts an Beobachtungen (Tabelle 11.2) und Unbekannten (Tabelle 11.3) soll in Tabelle 11.4 zur Ermittlung der maximalen Anzahl an Bedingungen fortgesetzt werden.

## 11.5 Funktionales Modell

Das Ziel des folgenden Abschnitts ist die Herstellung der linearisierten Form des funktionalen Modells (4.15) bzw. (4.25) in Matrixendarstellung. Die soeben in den Abschnitten 11.2 und 11.4 formu-

**Tab. 11.4:** Synthetisches Beispiel zur Ermittlung der maximalen Anzahl an Bedingungen.

Bedingung	$I$	$J_i$	$K_{ij}$	$b$	
Kugelbedingung	15	25	30	$\sum^I \sum^{J_i} K_{ij}$	11.250
Kollinearitätsbedingung	15	25	30	$2 \cdot \sum^I \sum^{J_i} K_{ij}$	22.500
Orthonormalitätsbedingung	15			$\sum^I \sum^{J_i} K_{ij}$	15
Datumsfestlegung					7
Gesamtanzahl der Bedingungen					33.772

lierten Beobachtungs- und Bedingungsgleichungen bilden entsprechend den theoretischen Ausführungen des [Abschnitts 4.4](#) die hierfür notwendige Grundlage. Im Folgenden werden die ursprünglichen und anschließend an der Stelle der Näherungswerte durch eine Reihenentwicklung nach Taylor linearisierten Verbesserungsgleichungen aufgestellt ([Abschnitte 11.5.1 bis 11.5.3](#)) und daraus die Koeffizientenmatrizen und Beobachtungs- bzw. Widerspruchsvektoren abgeleitet ([Abschnitte 11.5.4 und 11.5.5](#)).

### 11.5.1 Ursprüngliche und linearisierte Verbesserungsgleichungen

Basierend auf den aufgestellten Beobachtungsgleichungen (11.1) bzw. (11.2) ergeben sich die ursprünglichen und linearisierten Verbesserungsgleichungen der Bild- und Entfernungsmessungen zu

$$\begin{aligned} x'_{ij} + vx'_{ij} &= \Phi_{ij}^{Ax'}(X_{0i}, Y_{0i}, Z_{0i}, \omega_i, \varphi_i, \kappa_i, x'_0, c, \Delta x', X_j, Y_j, Z_j) \\ y'_{ij} + vy'_{ij} &= \Phi_{ij}^{Ay'}(X_{0i}, Y_{0i}, Z_{0i}, \omega_i, \varphi_i, \kappa_i, y'_0, c, \Delta y', X_j, Y_j, Z_j) \end{aligned} \quad (11.16)$$

$$\begin{aligned} \mathbf{v}\mathbf{x}'_{ij} &= \left( \frac{\partial \Phi_{ij}^A}{\partial \mathbf{X}_{0i}} \right)^0 d\mathbf{X}_{0i} + \left( \frac{\partial \Phi_{ij}^A}{\partial (\omega\varphi\kappa)_i} \right)^0 d(\omega\varphi\kappa)_i \\ &+ \left( \frac{\partial \Phi_{ij}^A}{\partial \mathbf{x}'_0} \right)^0 d\mathbf{x}'_0 + \left( \frac{\partial \Phi_{ij}^A}{\partial c} \right)^0 dc + \left( \frac{\partial \Phi_{ij}^A}{\partial \Delta \mathbf{x}'} \right)^0 d\Delta \mathbf{x}' + \left( \frac{\partial \Phi_{ij}^A}{\partial \mathbf{X}_j} \right)^0 d\mathbf{X}_j \\ &- (\mathbf{x}'_{ij} - \mathbf{x}'_{ij}{}^0) \end{aligned} \quad (11.17)$$

bzw.

$$D_{ijk} + vD_{ijk} = \Phi_{ijk}^D(X_{0i}, Y_{0i}, Z_{0i}, d_{[0,6]}, X_j, Y_j, Z_j) \quad (11.18)$$

$$\begin{aligned} vD_{ijk} &= \left( \frac{\partial \Phi_{ijk}^D}{\partial \mathbf{X}_{0i}} \right)^0 d\mathbf{X}_{0i} + \left( \frac{\partial \Phi_{ijk}^D}{\partial d_{[0,6]}} \right)^0 dd_{[0,6]} + \left( \frac{\partial \Phi_{ijk}^D}{\partial \mathbf{X}_j} \right)^0 d\mathbf{X}_j \\ &- (D_{ijk} - D_{ijk}{}^0) \end{aligned} \quad (11.19)$$

### 11.5.2 Ursprüngliche und linearisierte fiktive Verbesserungsgleichungen

Die Nebenbedingungen (11.9) und (11.10) fordern, dass die im Entfernungsbild lokalisierten Segmente der Kreise im Objektraum auf einer Kugel liegen und dass jeder Kugelpunkt entsprechend dem Modell der Zentralprojektion durch das Projektionszentrum auf dem Sensor abgebildet wird. Die Bedingungsgleichungen können entweder über ein erweitertes Normalgleichungssystem oder als fiktive Beobachtungen in eine vermittelnde Ausgleichung integriert werden (Abschnitt 4.4.3).

Eine strenge Integration beider Bedingungen ist aus numerischer Sicht nicht sinnvoll. Die Kugelbedingung kann bei einer strengen Einhaltung der Kollinearitätsbedingung Singularitäten im Gleichungssystem verursachen. Geometrisch bedeutet dies, dass ein Kreispunkt, zumeist während den ersten Iterationen der Ausgleichung, durch die strenge Einhaltung der Kollinearitätsbedingung außerhalb des aktuellen Kugelmodells projiziert wird. Es existiert demnach keine Lösung für die Kugelbedingung. Der vorliegende Fall erfordert somit eine flexible Gewichtung dieser zwei geometrischen Restriktionen in z. B. dem Rahmen einer Varianzkomponentenschätzung (VKS; Abschnitt 4.6). Die Bedingungsgleichungen  $\Gamma^{Kugel}$  und  $\Gamma^{Koll}$  werden hierfür in fiktive Beobachtungsgleichungen  $\Phi^{Kugel}$  und  $\Phi^{Koll}$  überführt. Die Bedingungsanzahl  $b$  verringert sich folglich, die Beobachtungsanzahl  $n$  erhöht sich aber in gleichem Maße, so dass die Redundanz  $f$  des gesamten Systems konstant bleibt.

Für die Kugelbedingungen lauten die hierfür notwendigen ursprünglichen und linearisierten fiktiven Beobachtungsgleichungen:

$$0 + vr_{jk} = \Phi_{jk}^{Kugel}(X_j, Y_j, Z_j, X_{jk}, Y_{jk}, Z_{jk}, r) \quad (11.20)$$

$$vr_{ijk} = \left( \frac{\partial \Phi_{jk}^{Kugel}}{\partial \mathbf{X}_j} \right)^0 d\mathbf{X}_j + \left( \frac{\partial \Phi_{jk}^{Kugel}}{\partial \mathbf{X}_{jk}} \right)^0 d\mathbf{X}_{jk} - \left[ (X_{jk} - X_j)^2 + (Y_{jk} - Y_j)^2 + (Z_{jk} - Z_j)^2 - r^2 \right] \quad (11.21)$$

Für die Kollinearitätsbedingungen lauten sie:

$$\begin{aligned} 0 + vx'_{ijk} &= \Phi_{ijk}^{Kollx'}(X_{0i}, Y_{0i}, Z_{0i}, \omega_i, \varphi_i, \kappa_i, x'_0, c, \Delta x', x'_{ijk}, X_{jk}, Y_{jk}, Z_{jk}) \\ 0 + vy'_{ijk} &= \Phi_{ijk}^{Kolly'}(X_{0i}, Y_{0i}, Z_{0i}, \omega_i, \varphi_i, \kappa_i, x'_0, c, \Delta y', y'_{ijk}, X_{jk}, Y_{jk}, Z_{jk}) \end{aligned} \quad (11.22)$$

$$\begin{aligned} \mathbf{v}\mathbf{x}'_{ijk} &= \left( \frac{\partial \Phi_{ijk}^{Koll}}{\partial \mathbf{X}_{0i}} \right)^0 d\mathbf{X}_{0i} + \left( \frac{\partial \Phi_{ijk}^{Koll}}{\partial (\omega\varphi\kappa)_i} \right)^0 d(\omega\varphi\kappa)_i \\ &+ \left( \frac{\partial \Phi_{ijk}^{Koll}}{\partial \mathbf{x}'_0} \right)^0 d\mathbf{x}'_0 + \left( \frac{\partial \Phi_{ijk}^{Koll}}{\partial c} \right)^0 dc + \left( \frac{\partial \Phi_{ijk}^{Koll}}{\partial \Delta \mathbf{x}'} \right)^0 d\Delta \mathbf{x}' + \left( \frac{\partial \Phi_{ijk}^{Koll}}{\partial \mathbf{X}_{jk}} \right)^0 d\mathbf{X}_{jk} \\ &- \Phi_{ijk}^{Koll}(\mathbf{X}_{0i}, (\omega\varphi\kappa)_i, \mathbf{x}'_0, c, \Delta \mathbf{x}', \mathbf{x}'_{ijk}, \mathbf{X}_{jk}) \end{aligned} \quad (11.23)$$

### 11.5.3 Linearisierte Bedingungsgleichungen

Die verbleibenden Bedingungsgleichungen (11.11) bis (11.14) verlangen zum einen die Orthonormalität der Rotationsmatrizen und legen zum anderen die Netzlagerung fest. Sie werden streng mit höchstem Zwang integriert.

Eine Reihenentwicklung der ursprünglichen Bedingungsgleichung (11.11) an der Stelle der Näherungswerte der Quaternionen ergibt den folgenden Ausdruck:

$$0 = 2q_{1i}dq_{1i} + 2q_{2i}dq_{2i} + 2q_{3i}dq_{3i} + 2q_{4i}dq_{4i} - (q_{1i}^2 + q_{2i}^2 + q_{3i}^2 + q_{4i}^2 - 1) \quad (11.24)$$

Die Bedingungsgleichungen (11.12) bis (11.14) zur Datumsfestlegung liegen bereits in linearisierter Form vor.

### 11.5.4 Koeffizientenmatrizen

Die partiellen Ableitungen der oben aufgestellten und sogleich linearisierten Verbesserungsgleichungen nach allen Unbekannten bilden die Koeffizienten  $\left(\frac{\partial \Phi}{\partial \mathbf{X}}\right)^0$  der Designmatrix  $\mathbf{A}$ . Die Zeilenanzahl entspricht der Anzahl aller realen und fiktiven Beobachtungen, die Spaltenanzahl der Anzahl aller unbekannt Parameter. Aus Gründen der Übersichtlichkeit wird sie wie folgt zerlegt<sup>5</sup>:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_A & \mathbf{A}_D & \mathbf{A}_{Kugel} & \mathbf{A}_{Koll} \\ (n_A, u) & (n_D, u) & (b_{Kugel}, u) & (b_{Koll}, u) \end{bmatrix}^T \quad (11.25)$$

mit

$$\begin{aligned} a_{ij}^A &= F \left[ \left(\frac{\partial \Phi_{ij}^A}{\partial \mathbf{X}_{0i}}\right)^0, \left(\frac{\partial \Phi_{ij}^A}{\partial (\omega \varphi \kappa)_i}\right)^0, \left(\frac{\partial \Phi_{ij}^A}{\partial \mathbf{x}'_0}\right)^0, \left(\frac{\partial \Phi_{ij}^A}{\partial c}\right)^0, \left(\frac{\partial \Phi_{ij}^A}{\partial \Delta \mathbf{x}'}\right)^0, \left(\frac{\partial \Phi_{ij}^A}{\partial \mathbf{X}_j}\right)^0 \right] \\ a_{ijk}^D &= F \left[ \left(\frac{\partial \Phi_{ijk}^D}{\partial \mathbf{X}_{0i}}\right)^0, \left(\frac{\partial \Phi_{ijk}^D}{\partial d_{[0,6]}}\right)^0, \left(\frac{\partial \Phi_{ijk}^D}{\partial \mathbf{X}_j}\right)^0 \right] \\ a_{jk}^{Kugel} &= F \left[ \left(\frac{\partial \Phi_{jk}^{Kugel}}{\partial \mathbf{X}_j}\right)^0, \left(\frac{\partial \Phi_{jk}^{Kugel}}{\partial \mathbf{X}_{jk}}\right)^0 \right] \\ a_{ijk}^{Koll} &= F \left[ \left(\frac{\partial \Phi_{ijk}^{Koll}}{\partial \mathbf{X}_{0i}}\right)^0, \left(\frac{\partial \Phi_{ijk}^{Koll}}{\partial (\omega \varphi \kappa)_i}\right)^0, \left(\frac{\partial \Phi_{ijk}^{Koll}}{\partial \mathbf{x}'_0}\right)^0, \left(\frac{\partial \Phi_{ijk}^{Koll}}{\partial c}\right)^0, \left(\frac{\partial \Phi_{ijk}^{Koll}}{\partial \Delta \mathbf{x}'}\right)^0, \left(\frac{\partial \Phi_{ijk}^{Koll}}{\partial \mathbf{X}_{jk}}\right)^0 \right] \end{aligned}$$

Zusätzlich wird die Koeffizientenmatrix  $\mathbf{B}$  benötigt, gebildet aus den partiellen Ableitungen  $\left(\frac{\partial \Gamma}{\partial \mathbf{X}}\right)^0$  der linearisierten Bedingungsgleichungen:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{Quat} & \mathbf{B}_{Dat} \\ (b_{Quat}, u) & (b_{Dat}, u) \end{bmatrix}^T \quad (11.26)$$

<sup>5</sup> Die Indizierung der einzelnen Elemente richtet sich in diesem und den folgenden Kapiteln nicht nach der diskreten Zeilen- bzw. Spaltenposition innerhalb der Matrix, sondern lediglich nach der bereits eingeführten Unterscheidung in *Aufnahme*, *Kugel* und *Kugeloberflächenpunkt*.

mit

$$b_i^{Quat} = F \left[ \left( \frac{\partial \Gamma_i^{Quat}}{\partial q_{1i}} \right)^0, \left( \frac{\partial \Gamma_i^{Quat}}{\partial q_{2i}} \right)^0, \left( \frac{\partial \Gamma_i^{Quat}}{\partial q_{3i}} \right)^0, \left( \frac{\partial \Gamma_i^{Quat}}{\partial q_{4i}} \right)^0 \right]$$

$$b_{[1,7]}^{Dat} = F \left[ \left( \frac{\partial \Gamma_{[1,7]}^{Dat}}{\partial \mathbf{X}_{0i}} \right)^0, \left( \frac{\partial \Gamma_{[1,7]}^{Dat}}{\partial \mathbf{X}_j} \right)^0, \left( \frac{\partial \Gamma_{[1,7]}^{Dat}}{\partial \mathbf{X}_{jk}} \right)^0 \right]$$

Die Koeffizientenmatrizen  $\mathbf{A}$  und  $\mathbf{B}$  enthalten zahlreiche Nullelemente. Diese werden vor allem durch die unbekanntes 3D-Kugeloberflächenkoordinaten hervorgerufen, welche keine direkte funktionale Verknüpfung zueinander aufweisen. Im Rahmen der Arbeit werden solche dünn- oder schwachbesetzten Matrizen durch sogenannte *Sparse*-Techniken effizient gespeichert und weiter numerisch behandelt (Press u. a. 1992).

### 11.5.5 Beobachtungs- und Widerspruchsvektoren

Der Beobachtungsvektor  $\mathbf{l}$  beinhaltet die Differenzen zwischen den gemessenen und den berechneten Beobachtungen (verkürzte Beobachtungen) mit folgender Struktur:

$$\mathbf{l}_{(n,1)} = \begin{bmatrix} \mathbf{l}_A & \mathbf{l}_D & \mathbf{l}_{Kugel} & \mathbf{l}_{Koll} \\ (n_A,1) & (n_D,1) & (b_{Kugel},1) & (b_{Koll},1) \end{bmatrix}^T \quad (11.27)$$

mit

$$l_{ij}^A = \mathbf{x}'_{ij} - \mathbf{x}_{ij}^0$$

$$l_{ijk}^D = D_{ijk} - D_{ijk}^0$$

$$l_{jk}^{Kugel} = \Gamma_{jk}^{Kugel}$$

$$l_{ijk}^{Koll} = \Gamma_{ijk}^{Koll}$$

Analog enthält der Vektor  $\mathbf{w}$  die Widersprüche:

$$\mathbf{w}_{(b,1)} = \begin{bmatrix} \mathbf{w}_{Quat} & \mathbf{w}_{Dat} \\ (b_{Quat},1) & (7,1) \end{bmatrix}^T \quad (11.28)$$

mit

$$w_i^{Quat} = \Gamma_i^{Quat}$$

$$w_{[1,7]}^{Dat} = 0$$

## 11.6 Stochastisches Modell

Nachdem im vorherigen [Abschnitt 11.5](#) die funktionalen Zusammenhänge zwischen den verschiedenen Beobachtungsarten einer 3D-Kamera sowie den zusätzlich formulierten geometrischen Nebenbedingungen und den zu bestimmenden unbekanntes (Kalibrier-)Parametern entwickelt wurden, beschäftigt sich dieser Abschnitt mit der Aufstellung des stochastischen Modells, also den stochastischen Eigenschaften der Beobachtungen und Nebenbedingungen (vgl. [Abschnitt 4.4.2](#)).

Das stochastische Modell wird vor der Ausgleichung durch die Kovarianzmatrix  $\mathbf{\Sigma}_{ll}$  der ursprünglichen Beobachtungen aufgestellt. Für eine adäquate stochastische Integration von Messwerten unterschiedlicher Art mit zueinander unbekanntem Genauigkeitsverhältnissen bietet sich die Bestimmung der zur Aufstellung von  $\mathbf{\Sigma}_{ll}$  notwendigen a-priori Varianzen im Rahmen einer VKS an.

Die integrierte Bündelblockausgleichung beinhaltet insgesamt  $g = 4$  Beobachtungsarten, nämlich

- Bildkoordinaten, gemessen im Amplitudenbild (Abschnitt 11.2.2),
- Distanzen, segmentiert aus dem Entfernungsbild (Abschnitt 11.2.3),
- Kugelbedingungen (Abschnitt 11.4.1) und
- Kollinearitätsbedingungen (Abschnitt 11.4.2),

für welche entsprechend dem Modell der VKS (Abschnitt 4.6.2) je eine Gruppenvarianz  $s_i^2$  eingeführt und die Kovarianzmatrix  $\mathbf{\Sigma}_{ll}$  wie folgt zerlegt wird:

$$\mathbf{\Sigma}_{ll} = \text{diag} \left( \begin{array}{cccc} \mathbf{\Sigma}_A & \mathbf{\Sigma}_D & \mathbf{\Sigma}_{Kugel} & \mathbf{\Sigma}_{Koll} \\ (n_A, n) & (n_D, n) & (b_{Kugel}, n) & (b_{Koll}, n) \end{array} \right) \quad (11.29)$$

mit

$$\begin{aligned} \sigma_{ll,ij}^A &= s_A^2 \\ \sigma_{ll,ijk}^D &= s_D^2 = s_{D_{konst}}^2 + s_{D_{linear}}^2 \cdot D \\ \sigma_{ll,jk}^{Kugel} &= s_{Kugel}^2 \\ \sigma_{ll,ijk}^{Koll} &= s_{Koll}^2 \end{aligned}$$

Die Zusammenfassung der einzelnen Bildkoordinaten- bzw. Distanzmessungen zu jeweils einer Beobachtungsklasse wird aus statistischer Sicht durch die Untersuchungen von R. Hempel und Westfeld (2009) bestätigt. Aus geodätischer Sicht ist eine Gruppierung sinnvoll, da alle Messungen einer Beobachtungsklasse dem selben Sensor entstammen und somit den selben zufälligen Fehlern unterliegen. Auch konnte für die Genauigkeit der Koordinatenmessungen in den Amplitudenbildern keine Abhängigkeit von der Entfernung erkannt werden (Abschnitt 11.2 und Abbildung 11.4), was die Zusammenfassung zu einer gemeinsam zu schätzenden Gruppenvarianz  $s_A^2$  ebenfalls rechtfertigt. Die Distanzmessgenauigkeit wird in dieser Arbeit durch einen einheitlichen Fehler  $s_D^2$  berücksichtigt (Abschnitt 9.5), kann aber prinzipiell auf zwei adaptive Varianzkomponenten in Form eines konstanten Fehleranteils  $s_{D_{konst}}^2$  und eines linear von der Distanzmessung selbst abhängigen Fehleranteils  $s_{D_{linear}}^2$  erweitert werden (vgl. Gleichung 11.29). Die interessante Frage, inwiefern dies die Genauigkeit des stochastischen Modells bzw. die Parameterschätzung signifikant verbessert, bleibt künftigen Untersuchungen vorbehalten.

Die notwendigen Näherungswerte der Gruppenvarianzen entstammen für die Bildkoordinatenmessungen den quadratischen Mittelwerten der a-posteriori LSM-Standardabweichungen ( $s_A = 1/25$  px; Abschnitt 11.2.2) und für die Distanzmessungen den Genauigkeitsinformationen durchgeführter Voruntersuchungen ( $s_D = 10$  mm; M. Hempel 2007). Darüber hinaus werden die Nebenbedingungen in Form fiktiver Beobachtungen empirisch mit  $s_{Kugel} = 1$  mm<sup>2</sup> und  $s_{Koll} = e^{-6}$  mm initialisiert.

Nach Gleichung (4.16) kann durch die Multiplikation der a-priori festgelegten Varianz  $s_0^2$  der Gewichtseinheit mit der soeben aufgestellten Kovarianzmatrix  $\Sigma_{ll}$  der Beobachtungen die Kofaktormatrix  $\mathbf{Q}_{ll}$  der Beobachtungen berechnet werden. Die Invertierung der Hauptdiagonalelemente  $q_{l,l_i}$  ergibt nach z. B. Gleichung (4.41) die zugehörigen Gewichte  $p_i$  einer Beobachtung. Die verbleibenden Nebenbedingungen (11.11) bis (11.14) werden im erweiterten Normalgleichungssystem (4.27) streng durch  $\mathbf{P}_B = \mathbf{0}$  berücksichtigt (Snow 2002).

Nach der Ausgleichung liegen mit den Kovarianzmatrizen  $\Sigma_{xx}$  der Unbekannten,  $\Sigma_{\hat{\gamma}\hat{\gamma}}$  der ausgeglichenen Beobachtungen und  $\Sigma_{vv}$  der Verbesserungen statistische Genauigkeitsmaße vor (Abschnitt 4.5.2), welche im Abschnitt 11.8 zur Beurteilung der Ausgleichungsergebnisse herangezogen werden.

## 11.7 Lösung der Ausgleichungsaufgabe

Dieser Abschnitt fasst, in Verbindung mit der in Abbildung 11.10 dargestellten Programmstruktur, den gesamten und vollständig automatischen Berechnungsablauf der entwickelten integrierten Bündelblockausgleichung zusammen.

### 11.7.1 Eingabe

Die funktionalen Eingabewerte sind die in beiden Kanälen gemessenen Beobachtungen:

- Koordinaten  $\mathbf{x}'_{ij}$  der Kugelmittelpunkte im Amplitudenbild
- Schrägdistanz  $D_{ijk}$  zu jedem Oberflächenpunkt der Kugel

Weiterhin gehen die Initialwerte  $\mathbf{X}^0$  aller zu schätzenden Modellparameter in das funktionale Modell der Ausgleichung ein, welche aus folgenden Quellen stammen (vgl. Abschnitt 11.3):

- Herstellerangaben ( $c$ ,  $\mathbf{x}'_0$ ,  $\Delta\mathbf{x}'$ )
- Automatische Näherungswertberechnung ( $\mathbf{X}_{0i}$ ,  $(\omega\phi\kappa)_i$ ,  $d_{[0,6]}$ ,  $\mathbf{X}_{jk}$ )
- Herstellungsprozess des Referenzfelds ( $\mathbf{X}_j$ )

Darüber hinaus sind die folgenden stochastischen Eingabewerte bereitzustellen:

- A-priori Standardabweichung  $s_0$  der Gewichtseinheit
- A-priori Standardabweichungen  $s_A$ ,  $s_D$ ,  $s_{Kugel}$  und  $s_{Koll}$  der ursprünglichen (fiktiven) Beobachtungen

### 11.7.2 Modellaufstellung

Die der integrierten Bündelblockausgleichung übergebenen Eingabewerte werden verwendet, um das funktionale und stochastische Modell aufzustellen, welche in den Abschnitten 11.5 und 11.6 definiert sind. Für jede Iteration  $k$  der Ausgleichung werden die Koeffizientenmatrizen  $\mathbf{A}$  und  $\mathbf{B}$  sowie

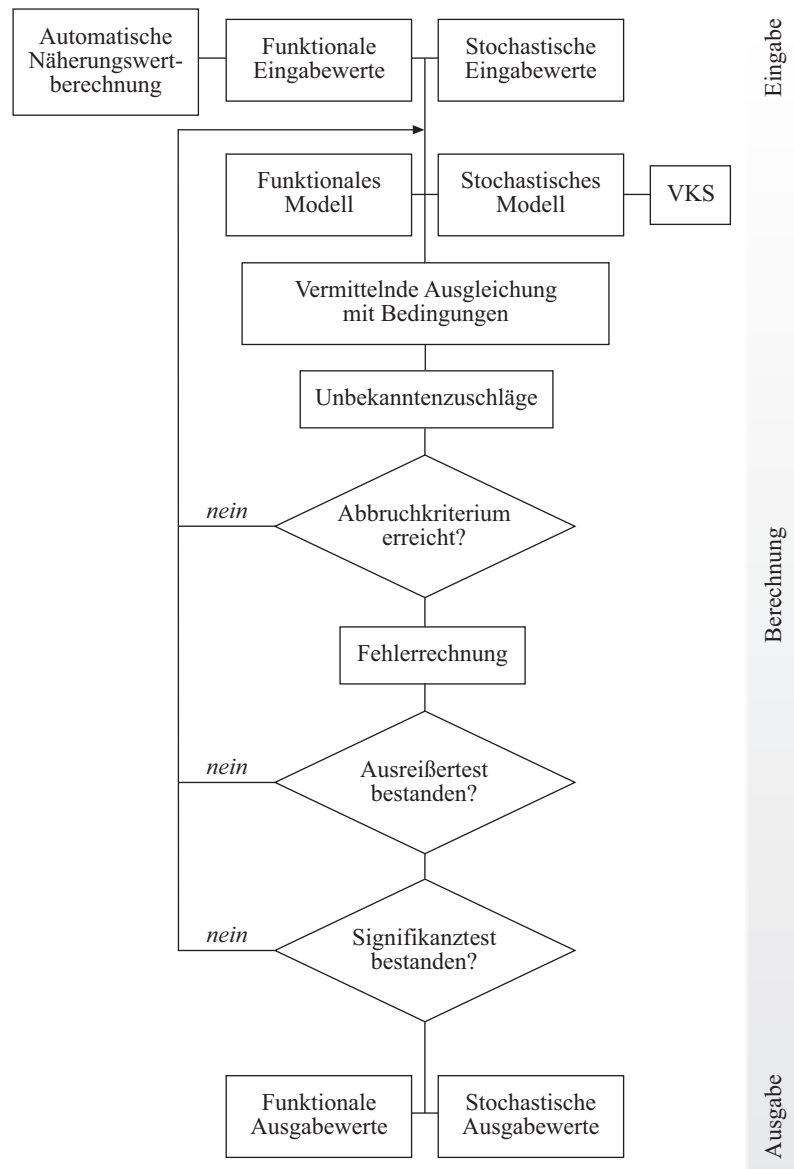


Abb. 11.10: Programmstruktur der integrierten Bündelblockausgleichung.

die Beobachtungs- bzw. Widerspruchsvektoren  $\mathbf{l}$  und  $\mathbf{w}$  neu, basierend auf dem aktuellen Modellparametersatz  $\mathbf{X}^k$ , gebildet. Ebenso werden in jeder Iteration die Varianzen  $s_j^2$  der ursprünglichen Beobachtungen im Rahmen einer VKS laut dem alternativen Berechnungsschema nach (Förstner 1979) geschätzt und zur Aufstellung der Gewichtsmatrix  $\mathbf{P}$  verwendet. Für die letzte Iteration wird auf den Kochschen Ansatz (2004) zurückgegriffen, um auch die Varianzen der Varianzkomponenten berechnen zu können.



### 11.7.3 MKQ

An dieser Stufe des Programms stehen alle für die Lösung der Ausgleichungsaufgabe notwendigen Informationen zu Verfügung. Entsprechend einer Ausgleichung nach vermittelnden Beobachtungen mit Bedingungen zwischen den Unbekannten ([Abschnitt 4.4.3](#)) wird das erweiterte Normalgleichungssystem ([4.27](#)) aufgestellt und nach dem Prinzip der **MKQ** gelöst.

### 11.7.4 Abbruchkriterium

Nach jeder Iteration  $k$  werden solange kleine geschätzte Zuschläge  $\hat{\mathbf{x}}$  an den Parametervektor  $\mathbf{X}^k$  angebracht ([Gleichung 4.21](#)), bis das folgende Abbruchkriterium erreicht ist:

$$1 - 1e^{-5} < \frac{(\mathbf{v}^T \mathbf{P} \mathbf{v})_k}{(\mathbf{v}^T \mathbf{P} \mathbf{v})_{k-1}} < 1 + e^{-5} \quad (11.30)$$

Konvergiert das Ausgleichungssystem nach einer zuvor definierten maximalen Anzahl an Iterationen nicht, findet ebenfalls ein Iterationsabbruch statt. Auch wird eine mögliche Oszillation des Lösungsvektors abgefangen.

### 11.7.5 Fehlerrechnung

Die folgende Fehlerrechnung nach dem in [Abschnitt 4.5](#) gegebenen Berechnungsapparat liefert aus den Kovarianzmatrizen der Beobachtungen, Unbekannten und Verbesserungen alle für eine statistische Auswertung notwendigen Qualitätsmaße. Hierzu gehören:

- A-posteriori Standardabweichung  $\hat{\sigma}_0$  der Gewichtseinheit
- A-posteriori Standardabweichungen  $\hat{\sigma}_l$  bzw.  $\hat{\sigma}_\hat{\gamma}$  der ursprünglichen bzw. ausgeglichenen Beobachtungen
- A-posteriori Standardabweichungen  $\hat{\sigma}_{\hat{\mathbf{x}}}$  der geschätzten unbekanntem Modellparameter
- A-posteriori Standardabweichungen  $\hat{\sigma}_v$  der Verbesserungen
- Koeffizienten  $\rho_{ij}$  zur Analyse möglicher Korrelationen zwischen den geschätzten Parametern
- Die Elemente der Matrix  $\mathbf{D}$  als Maße für die Bestimmbarkeit eines Parameters

### 11.7.6 Ausreißertest

Die im Zuge der Fehlerrechnung ausgegebene Standardabweichung  $\hat{\sigma}_{v_i}$  wird ins Verhältnis zur eigentlichen Verbesserung  $v_i$  gesetzt. Der Quotient ([4.37](#)) ist die normierte Verbesserung  $w_i$ , welche nach dem Vorbild des in [Abschnitt 4.5.4](#) eingeführten *Data Snooping*s als ein Zuverlässigkeitsmaß zur Detektion von Ausreißern im Beobachtungsmaterial dient. Grobe Fehler mit einer normierten Verbesserung  $w_i$  größer als der zuvor definierte **GWDS**  $\varepsilon_{GWDS} = 2,5$  sind auszuschließen. Für die Bildkoordinatenmessungen wird stets nur der größte Ausreißer eliminiert, da dieser die Verbesserungen von eigentlich fehlerfreien Beobachtungen beeinflussen kann. Gleiches gilt prinzipiell auch für die Distanzmessungen als zweite Beobachtungsgruppe, jedoch kann dies aufgrund der großen

Anzahl an Beobachtungen im Sinne einer vernünftigen Balance zwischen Rechenzeit und Systemredundanz vernachlässigt werden.

### 11.7.7 Signifikanztest

Nachdem eine Modelllösung aus grob fehlerfreien Beobachtungen geschätzt wurde, kann die Signifikanz eines Parameters im funktionalen Modell geprüft werden ([Abschnitt 4.5.6](#)). In der vorliegenden Implementierung wird das Quantil  $t_{p,f}$  der t-Verteilung für einen zweiseitigen Vertrauensbereich über die Wahrscheinlichkeit  $p = 0,95$  sowie die Anzahl an Freiheitsgraden  $f$  berechnet ([Abramowitz & Stegun 1964](#)). Ist der Quotient  $\hat{x}_i/\hat{s}_{\hat{x}_i}$  kleiner als  $t_{p,f}$ , wird der Parameter eliminiert und die Ausgleichung erneut durchlaufen.

### 11.7.8 Ausgabe

Die hauptsächlich interessierenden Ausgabewerte sind:

- Geschätzte Modellparameter  $\hat{x}_i$  mit ihren a-posteriori Standardabweichungen  $\hat{s}_{\hat{x}_i}$
- Residuen  $v_i$  der Beobachtungen
- Genauigkeits- und Zuverlässigkeitsmaße, berechnet im Zuge der oben genannten **VKS** und Fehlerrechnung

## 11.8 Validierung der integrierten Bündelblockausgleichung

### 11.8.1 Experimentalkonfiguration

#### Hardware

Die entwickelte integrierte Bündelblockausgleichung wird zur Kalibrierung von zwei baugleichen 3D-Kameras PMD[vision]<sup>®</sup> CamCube 2.0 und einer 3D-Kamera SwissRanger<sup>™</sup> SR4000 verwendet. Die folgenden Untersuchungen zur Validierung des Verfahrens beziehen sich auf die Daten und Ergebnisse eines der beiden Modelle eines PMD[vision]<sup>®</sup> CamCube 2.0 ([Abbildung 11.11](#)). Eine vergleichende Analyse von 3D-Kameras derselben Art bzw. unterschiedlicher Hersteller ist prinzipiell sehr interessant, wurde aber im Rahmen dieser Arbeit nicht realisiert.

Distanzmessende Kameras ermöglichen dem Nutzer die Variationen von zwei wesentlichen Aufnahmeparametern, nämlich der Modulationsfrequenz  $f$  und der Integrationszeit  $\Delta t$  ([Abschnitt 8.2](#)). In dem hier dokumentierten Experiment beträgt  $f = 20$  MHz, was einem maximal eindeutigen Distanzmessbereich von  $D_{max} = 7,5$  m entspricht. Die Integrationszeit wird mit  $\Delta t = 2.000 \mu s$  auf einen für die im [Teil IV](#) vorgestellten bewegungsanalytischen Anwendungen zweckmäßigen Wert festgesetzt. Da bei unterschiedlichen Modulationsfrequenzen und Integrationszeiten unterschiedliche Größenordnungen der auf eine 3D-Kamera wirkenden Fehlereinflüsse zu erwarten sind ([Kapitel 9](#)), ist zukünftig die Datenaufnahme und -prozessierung unter der Verwendung der für die jeweilige Anwendung adäquaten Aufnahmeparameter sinnvoll.

## Referenzfeld

Die Wahl der Referenzfeldgröße unterliegt mehreren in (Meißner 2010) detailliert beschriebenen Einflussfaktoren. Eine große Dimensionierung erlaubt die Einhaltung eines ausreichenden Abstands zwischen den einzelnen Referenzkugeln, was sowohl für die spätere Kugelsegmentierung (Abschnitt 11.2.3) als auch zur Reduktion von Mehrwegeeffekten zwischen den einzelnen Kugeln zweckmäßig ist (Abschnitt 9.3.1). Weiterhin können so, insbesondere in größeren Entfernungen, formatfüllende Abbildungen und folglich die sichere Bestimmung der äußeren Kameraorientierungen gewährleistet werden. Dem gegenüber steht die aufgestellte Forderung nach einer möglichst kompakten Bauweise. Mit einem Grundriss von  $(100\text{cm})^2$  wird ein Kompromiss eingegangen (Abbildung 11.11).

Für die Herstellung des Referenzfelds finden 25 weiße Styroporkugeln mit einem Radius  $R$  von 35mm Verwendung. Auch hier ist die Wahl des Radius ein Kompromiss zwischen der Größe der Kalibrierplatte und der Größe der Kugelprojektion, also der Anzahl der pro Kugel abgebildeten Oberflächenpunkte als Funktion der Entfernung (Meißner 2010). In der vorliegenden Konfiguration sind die Kugeln in zwei Ebenen über 50cm bzw. 75cm lange Gewindestangen auf der Grundplatte fixiert.

Weiterhin sei konstatiert, dass, aufgrund der Integration zusätzlicher Entfernungsmessungen, ein dreidimensionales Testfeld zur stabilen Rekonstruktion einer 3D-Kameraorientierung im Bündel prinzipiell nicht notwendig ist. Aufgrund der in Abschnitt 11.1 beschriebenen Grundidee einer direkten funktionalen Verknüpfung von im Amplituden- und Entfernungskanal beobachteten Größen, ist der Einsatz eines ebenen Referenzfeld jedoch nicht möglich.

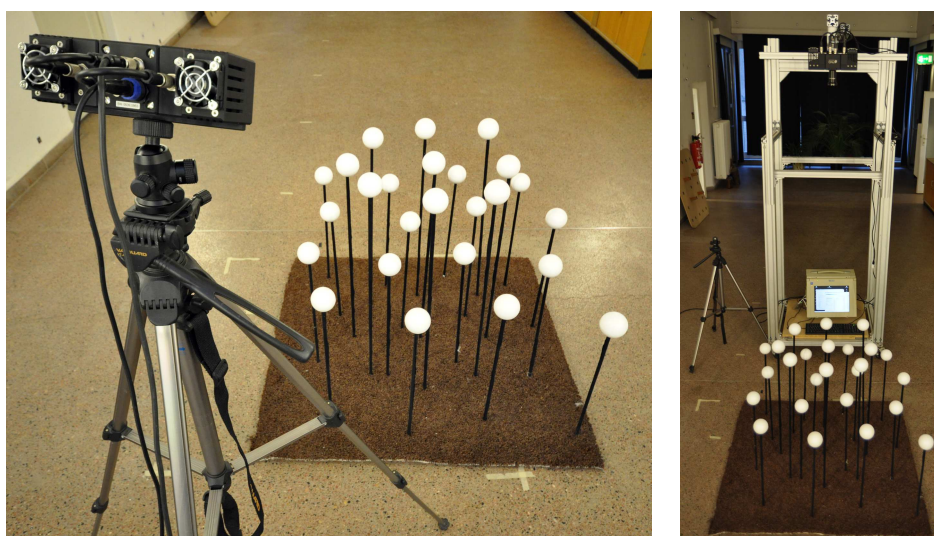


Abb. 11.11: Referenzfeld.

## Datenaufnahme

Für die Kalibrierung wird die in Abbildung 11.12 schematisch dargestellte Netzkonfiguration vorgeschlagen. Insgesamt zwölf Konvergentaufnahmen, davon vier jeweils um  $90^\circ$  um die Aufnahmeach-

se gekantete Frontalaufnahmen, stellen eine optimale Strahlenschnittgeometrie sicher und vermeiden unerwünschte Korrelationen zwischen den Orientierungsparametern (Abschnitt 5.4). Weitere vier Kamerastandpunkte in einem Abstand von bis zu 5,5 m ermöglichen die Bestimmung distanzabhängiger Korrekturterme für einen möglichst großen Messbereich (Abschnitt 11.3.3). Zusätzliche Aufnahmen in weiteren Entfernungen sind prinzipiell problemlos in die Ausgleichung integrierbar, jedoch bei der gegebenen Experimentalkonfiguration nicht zuverlässig zu orientieren.

Die Aufnahme einer 3D-Kamera wird durch eine Vielzahl von Rauschquellen beeinflusst, wie z. B. Hintergrundstrahlung, thermisches Rauschen, *Fixed Pattern Noise* (FPN), Quantenrauschen oder Dunkelstromrauschen (Kapitel 9; vgl. z. B. auch Lange 2000). Im Vergleich zu einer Einzelaufnahme kann nicht modellierbares Rauschen durch  $N$  Wiederholungsmessungen auf  $(1/\sqrt{N})$  reduziert werden. Folglich werden bei dem beschriebenen Experiment pro Standpunkt 100 Aufnahmen erzeugt und anschließend arithmetisch gemittelt.

### 11.8.2 Reduktion der Distanzbeobachtungen

Bei der in Abbildung 11.12 gezeigten Netzkonfiguration kommt es aufgrund der für die sichere Bestimmung der Kameraorientierungsparameter erforderlichen Schräg- und gewälzten Frontalaufnahmen (Abschnitt 5.4.3) zu einer stark ungleichmäßigen Verteilung der Distanzbeobachtungen zu Gunsten des vorderen Drittels des gesamten Messbereichs. Aus diesem Grund werden lediglich die Distanzbeobachtungen der Aufnahmen 13 bis 16 verwendet.

Weiterhin erfolgt eine Ausdünnung der Distanzmessungen aller in Abschnitt 11.2.3 vollständig segmentierten Kugeloberflächenpunkte. Dies ist zum einen notwendig, um die an den Kugelrändern auftretenden Kometenschweife zu eliminieren (Gordon 2008), zum anderen, um eine Gleichverteilung der Messungen zu nahen und weit entfernten Kugeln zu gewährleisten. Darüber hinaus verlangt die implementierte VKS eine Gleichverteilung der Beobachtungsanzahl der einzelnen Beobachtungsgruppen.

Eine genaue Quantifizierung des Einflusses der Ausdünnung bleibt zukünftigen Arbeiten vorbehalten.

### 11.8.3 Innere Orientierung und Zusatzparameter

Die aus der integrierten Bündelblockausgleichung resultierenden Elemente  $\hat{x}_i$  der inneren Orientierung und deren a-posteriori Standardabweichungen  $\hat{s}_{\hat{x}_i}$  sind für ein 3D-Kameramodell PMD[vision]<sup>®</sup> CamCube 2.0 in Tabelle 11.5 gelistet. Bis auf den Verzeichnungsparameter  $A_3$  sowie die Affinitäts- und Scherungseinflüsse  $C_{1,2}$  sind alle Parameter signifikant bestimmt worden. Die a-posteriori Standardabweichung  $\hat{s}_c$  der Kamerakonstante  $c$  liegt mit  $4,93 \mu\text{m}$  (ca.  $1/10\text{px}$ ) zwar über der Größenordnung der erreichbaren Bildmessgenauigkeit, bewirkt im Objektraum dennoch einen für dieses Messgerät realistischen Maßstabsfehler von nur etwa  $0,4\%$  der Objektgröße. Die Verschiebung der Hauptpunktlage bewegt sich mit ca.  $4\text{px}$  in  $-x'$ - und ca.  $5\text{px}$  in  $-y'$ -Richtung in einem realistischen

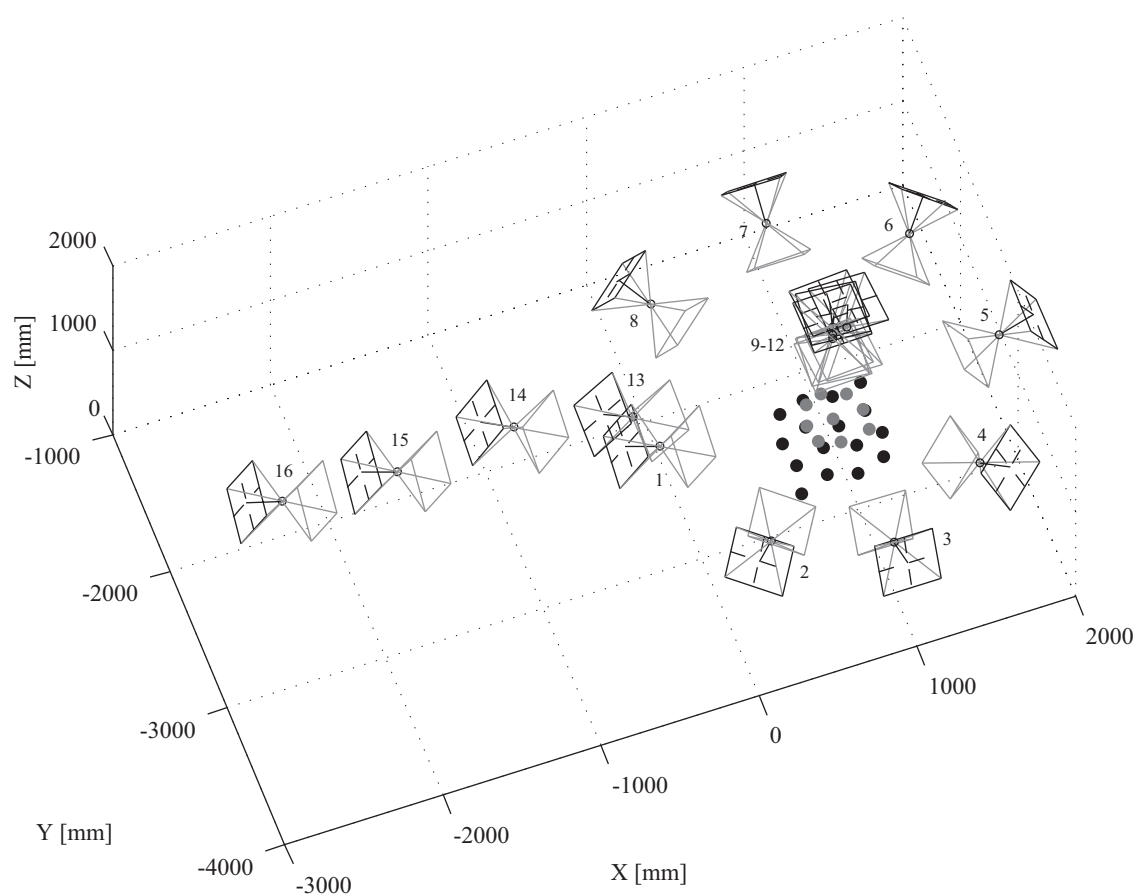


Abb. 11.12: Beispiel für die Netzkonfiguration eines 3D-Kamerabildverbands.

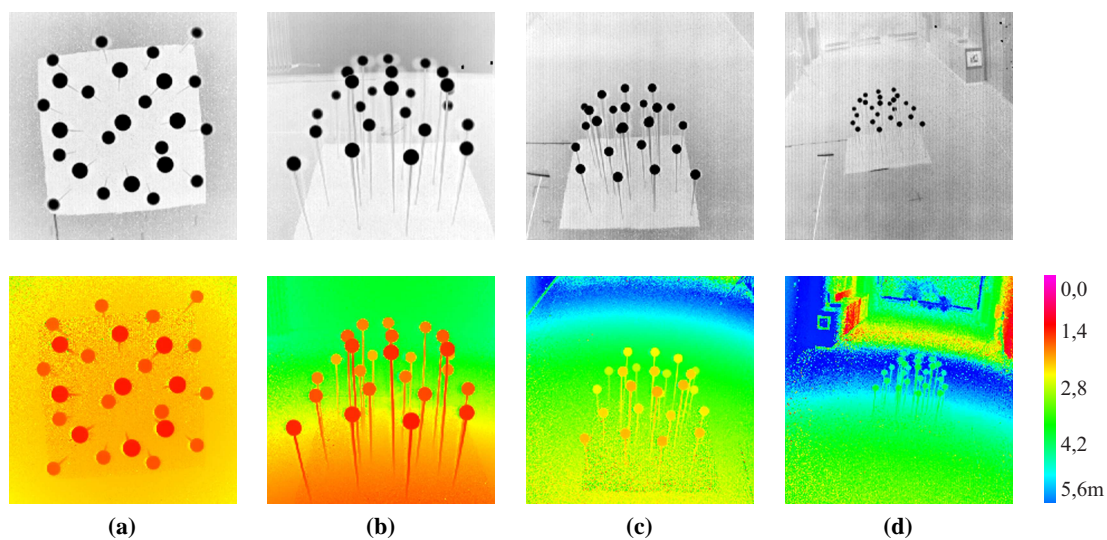


Abb. 11.13: Inverse Amplitudenbild- (oben) und farbkodierte Entfernungsbildaufnahmen (unten) zur integrierten Bündelblockausgleichung. (a) Frontalaufnahme. (b – d) Schrägaufnahmen im Abstand von ca. 1,5 m bis 4,5 m.

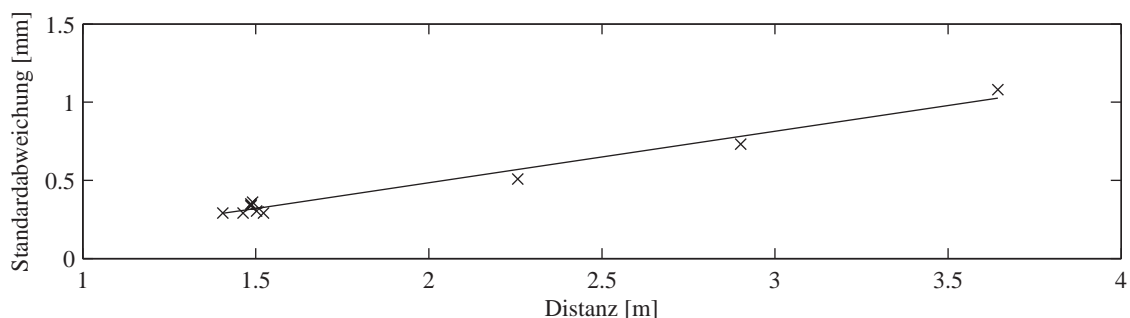
Bereich. Die a-posteriori Standardabweichungen  $\hat{\delta}_{A_{1,2}, B_{1,2}}$  der signifikant geschätzten Verzeichnungsparameter  $A_{1,2}$  und  $B_{1,2}$  sind deutlich niedriger als die Parameter selbst.

**Tab. 11.5:** Aus der integrierten Bündelblockausgleichung resultierende Parameter  $\hat{x}_i$  der inneren Orientierung und von Bildfehler beschreibenden Funktionen sowie deren a-posteriori Standardabweichungen  $\hat{\delta}_{\hat{x}_i}$  einer 3D-Kamera PMD[vision]<sup>®</sup> CamCube 2.0.

	$c$ [mm]	$x'_0$ [mm]	$y'_0$ [mm]	$A_1$	$A_2$	$A_3$	$B_1$	$B_2$	$C_1$	$C_2$
$\hat{x}_i$	12,179	-0,176	-0,230	$3,16e^{-3}$	$1,37e^{-5}$	0	$4,13e^{-4}$	$6,54e^{-4}$	0	0
$\hat{\delta}_{\hat{x}_i}$	$4,93e^{-3}$	$7,66e^{-3}$	$7,00e^{-3}$	$3,20e^{-5}$	$1,31e^{-6}$	fix	$1,98e^{-5}$	$1,89e^{-5}$	fix	fix

#### 11.8.4 Äußere Orientierung

Die Parameter der äußeren Orientierung wurden eindeutig bestimmt und sollen nicht weiter betrachtet werden. Einzig die in [Abbildung 11.14](#) gegen die mittleren Aufnahmedistanzen dargestellten **RMS-Abweichungen** der a-posteriori Standardabweichungen  $\hat{\delta}_{\mathbf{X}_0}$  der Projektionszentren  $\mathbf{X}_0$  verdeutlichen, dass der zu kalibrierende Messbereich auch aufgrund der geforderten Abbildungsgeometrie eingeschränkt ist. Das durch die Kugelmittelpunkte  $\mathbf{X}_j$  und das jeweilige Projektionszentrum  $\mathbf{X}_{0i}$  aufgespannte Strahlenbündel lässt sich nur dann genau und zuverlässig in die Bildpunkte  $\mathbf{x}'_{ij}$  einpassen, wenn diese über die gesamte Bildebene verteilt sind. In zu großen Aufnahmeentfernungen verringert sich nicht nur der abgebildete Kreisdurchmesser und folglich die zur Auswertung verwendbaren Entfernungsbeobachtungen ([Meißner 2010](#)), es wird auch das Referenzfeld nicht mehr entsprechend den Forderungen des [Abschnitts 5.3.3](#) formatfüllend abgebildet ([Abbildung 11.13d](#)). Aufnahmen in einer Entfernung von mehr als 4,0m sind mit der gewählten Referenzfeldgröße nicht mehr genau und zuverlässig zu orientieren. Da die Referenzfeldgröße und somit auch der zu kalibrierende Messbereich entsprechend der im Rahmen dieser Arbeit behandelten bewegungsanalytischen Aufgabenstellung ausgelegt wurde, ist dies als wenig kritisch anzusehen.



**Abb. 11.14:** Quadratisches Mittel  $RMS_{\mathbf{X}_0}$  der a-posteriori Standardabweichung  $\hat{\delta}_{\mathbf{X}_0}$  der Projektionszentren  $\mathbf{X}_{0i}$  in Abhängigkeit der mittleren Aufnahmedistanz  $\bar{D}_\perp$ .



### 11.8.5 Distanzkorrektur

Die Terme  $d_{[0,6]}$  zur Korrektur distanzspezifischer Fehlereinflüsse sind mit ihren a-posteriori Standardabweichungen in der [Tabelle 11.6](#) gelistet. Es konnten alle Parameter innerhalb des in [Abschnitt 11.7.7](#) angegebenen Signifikanzniveaus von 95 % eindeutig aus dem vorliegenden Beobachtungsmaterial in dem aufgestellten funktionalen Modell geschätzt werden.

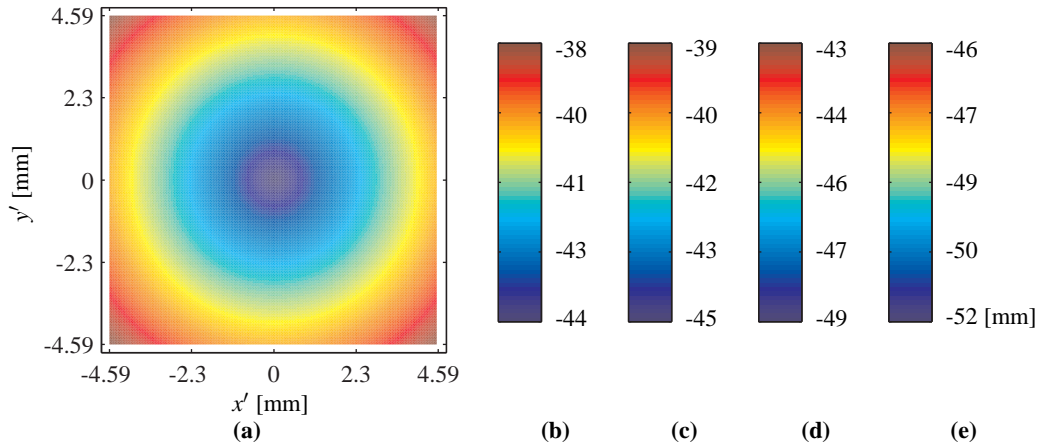
Der additive Term  $d_0$  des linearen Anteils  $\Delta D_{lin}$  liegt bei ca. 80 mm; der multiplikative Parameter  $d_1$  zur Modellierung des entfernungsabhängigen linearen Trends bewegt sich im Bereich von ca. 1,2 % der gemessenen Strecke. Beide Parameter entsprechen in ihren Größenordnung den aus Voruntersuchungen stammenden und in [Abschnitt 9.5.1](#) angegebenen Werten eines vergleichbaren Vorgängermodells SwissRanger<sup>TM</sup> SR3k. Die Amplituden  $d_2$  bis  $d_5$  der periodischen Korrekturfunktion  $\Delta D_{zykl}$  approximieren den zyklischen Phasenfehler und verursachen eine Auslenkung der Wellen um maximal  $\pm 30$  mm. Im Vergleich zu den in [Abbildung 9.3](#) dargestellten periodischen Soll-Ist-Abweichungen werden sie bei den 3D-Kameramodellen der neueren Generation vermutlich bereits vom Hersteller durch implementierte geräteinterne Korrekturmodelle deutlich reduziert. Der Radialfaktor  $d_6$  modelliert den sensorpositionsabhängigen Fehlereinfluss  $\Delta D_{x,y}$  und bewirkt Korrekturen von bis zu 6 mm für die korrespondierenden Distanzmesswerte der Bildrandpixel. Die [Abbildung 11.15](#) fasst die Auswirkungen der Distanzkorrektur  $\Delta D$  in Abhängigkeit von der Sensorposition und der Entfernung zusammen. Die in ([Pattinson 2010b](#)) quantifizierten Korrekturwerte entstammen der 2-Stufen-Kalibrierung eines PMD[vision]<sup>®</sup> CamCube 2.0 und sind in ihren Größenordnungen mit den hier erzielten Werten vergleichbar (vgl. [Abschnitt 10.2](#)). Einzig der radiale Einfluss ist in [Pattinsons](#) Beitrag deutlich größer.

Die einzelnen Standardabweichungen  $\hat{\sigma}_{\hat{x}_i}$  der Kalibrierparameter des Distanzkorrekturmodells erscheinen recht hoch und können aufgrund verbleibender und nicht berücksichtigter, da noch nicht parametrisierter Restsystematiken sehr schwer interpretiert werden ([Kapitel 9](#)). Eine Reduktion des Parametersatzes um bspw. die zyklischen und signalwegeffektbedingten Fehleranteile kann zwar die innere Genauigkeit bei der Bestimmung der verbleibenden zwei linearen Fehleranteile um eine Größenordnung deutlich verbessern ( $d_0 = -45,68 \text{ mm} \pm 0,989 \text{ mm}$ ;  $d_1 = 1,06e^{-2} \pm 7,54e^{-4}$ ), die Residuen der Distanzbeobachtungen im Gegenzug aber wiederum um ca. einen halben Zentimeter verschlechtern. Die nicht vollständig adäquate Modellierung distanzmessspezifischer Fehler äußert sich auch in den nach Gleichung (4.38) berechneten Korrelationskoeffizienten  $\rho_{ij}$ . Sämtliche Distanzkorrekturparameter weisen gegenseitige Korrelationen mit Koeffizienten größer 0,75 auf.

Die Kalibrierstrategien der Arbeitsgruppen um [Böhm und Pattinson \(2010\)](#) sowie [Lichti u. a. \(2010\)](#) sind in ihren Parametrisierungen mit dem vorliegenden Ansatz vergleichbar ([Kapitel 10](#)). In ihren Beiträgen liefern sie jedoch keine Genauigkeitsangaben, welche zur Beurteilung der hier vorliegenden Ergebnisse herangezogen werden können. An dieser Stelle kann somit nur auf die Aussagekraft der im [Abschnitt 11.8.7](#) folgenden Residuenanalyse verwiesen werden. Die in ([Kahlmann 2007](#)) thematisierte Fragestellung zur zeitlichen Stabilität der geschätzten Kalibrierparameter eines SwissRanger<sup>TM</sup> SR3k soll Gegenstand zukünftiger Untersuchungen sein.

**Tab. 11.6:** Parameter der Distanzkorrekturterme und ihre a-posteriori Standardabweichungen.

	$d_0$ [mm]	$d_1$	$d_2$ [mm]	$d_3$ [mm]	$d_4$ [mm]	$d_5$ [mm]	$d_6$
$\hat{x}_i$	-77,33	$1,20e^{-2}$	-19,94	10,67	-7,21	-5,42	0,924
$\hat{s}_{\hat{x}_i}$	9,83	$3,76e^{-3}$	3,90	1,71	1,10	1,17	0,432

**Abb. 11.15:** Auswirkungen der Distanzkorrektur  $\Delta D$ : Aus (a) sind deutlich die sensorpositionsabhängigen Fehleranteile ersichtlich. (b) bis (e) geben die zugehörigen farbkodierten Wertebereiche der Distanzkorrektur für Entfernungen von  $D_{\perp} = [1, 0; 4, 0]$  m an.

### 11.8.6 3D-Objektpunktkoordinaten

Zur Bewertung der Genauigkeit der geschätzten Neupunktkoordinaten, also der Koordinaten  $\mathbf{X}_j$  der Kugelmittelpunkte und  $\mathbf{X}_{jk}$  der Kugeloberflächenpunkte, können die quadratischen Mittelwerte der a-posteriori Standardabweichungen verwendet werden. Aufgrund der konvergenten Aufnahmeconfiguration entsprechend der [Abbildung 11.12](#) ist eine getrennte Betrachtung der drei Koordinatenrichtungen und die damit einhergehende Unterscheidung in Lateral- und Tiefenrichtung wenig sinnvoll. Vielmehr ist der Betrag  $RMS_{\mathbf{X}}$  des Fehlervektors zur Beurteilung heranzuziehen. Er beträgt für die Neupunktkoordinaten der Kugelmittelpunkte ca. 0,13 mm. Die Unsicherheiten bei der Bestimmung der Koordinaten der Kugeloberflächenpunkte fallen mit 0,32 mm geringfügig größer aus. Eine mögliche Begründung hierfür liegt in dem aufgestellten funktionalen Zusammenhang: Die Schätzung der Kugelmittelpunktkoordinaten basiert auf der räumlichen Einpassung von Strahlenbündeln unter der Forderung, dass sich aus mehreren Bildern rekonstruierte homologe Bildstrahlen optimal in ihrem Objektpunkt schneiden ([Abschnitt 5.4](#)). Die Schätzung der Koordinaten der Kugeloberflächenpunkte basiert aufgrund der indirekten Integration des Kugelmittelpunkts in die Kugelbedingung zwar ebenfalls auf diesem geometrischen Zusammenhang, verwendet aber direkt Bildkoordinaten- und Entfernungsmessungen, welche von lediglich einem Standpunkt aus beobachtet wurden ([Abschnitt 11.1](#)).

Die aus den durchschnittlichen a-posteriori Standardabweichungen der geschätzten Neupunktko-



ordinaten gerechneten **RMS**-Abweichungen spiegeln lediglich die innere Genauigkeit wider. Ein Soll-Ist-Vergleich mit den aus den Daten eines Messsystems höherer Genauigkeit berechneten Koordinaten zur Ermittlung der äußeren Genauigkeit wird in **Abschnitt 11.8.9** gegeben.

### 11.8.7 Residuenanalyse

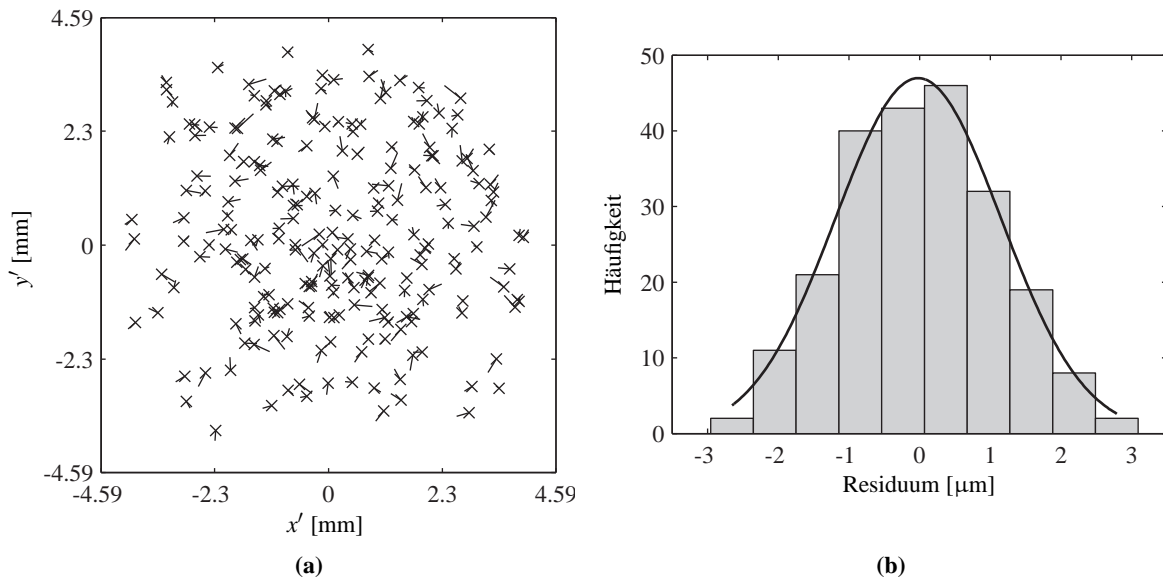
Die zufällig verteilten Restklaffungen  $v_A$  der Bildkoordinatenmessungen  $\mathbf{x}'_{ij}$  der Kugelmittelpunkte in den Amplitudenbildern folgen einer Normalverteilung und streuen in beiden Koordinatenrichtungen im quadratischen Mittel ihrer a-posteriori Standardabweichung  $\hat{s}_{v_A}$  von ca.  $1/40$  px um den Erwartungswert  $\mu = -0,02 \mu\text{m}$ . Die **Abbildung 11.16a** zeigt deren Beträge und Richtungen für die Messung der Bildkoordinaten aller Amplitudenbilder in einer gemeinsamen Darstellung; die **Abbildung 11.16b** veranschaulicht die zugehörige Häufigkeitsverteilung.

Die verbleibenden Residuen  $v_D$  der Distanzbeobachtungen sind in **Abbildung 11.17** vergleichend sowohl für den Ausgleichungsfall ohne als auch unter Berücksichtigung des aufgestellten Distanzkorrekturmodells dargestellt. Der Verlauf der in die Residuen der unkorrigierten Distanzmessungen eingepassten Funktion zeigt einen konstanten *Offset*, einen leicht entfernungsabhängigen linearen Verlauf sowie periodische Schwankungen. Deren Größenordnungen wurden bereits im **Abschnitt 11.8.5** vorgestellt und interpretiert. Das quadratische Mittel  $RMS_{v_D}$  der Residuen liegt mit  $40,86 \text{ mm}$  auf einem erwartungsgemäß hohen Niveau. Die Schätzung und Anbringung der in **Kapitel 9** ausführlich diskutierten Distanzkorrekturterme zur Modellierung linearer, zyklischer und signalwegeffektbedingter Distanzmessfehler reduziert die Abweichungen deutlich. Die in die verbleibenden Residuen ohne interpretierbare systematische Effekte eingepasste Funktion folgt nunmehr annähernd dem Verlauf einer Geraden mit  $y = 0 = \text{const}$  und bestätigt die in **Abbildung 11.17b** gezeigte Normalverteilung der Restklaffungen um den Erwartungswert  $\mu = 0,58 \text{ mm}$ . Die **RMS**-Abweichung konnte um den Faktor 5 auf  $8,66 \text{ mm}$  verringert werden, was sich mit den Untersuchungsergebnissen von **Böhm und Pattinson (2010)** sowie **Lichti u. a. (2010)** deckt.

### 11.8.8 Fehler der Beobachtungsgruppen

Die durch eine **VKS** automatisch mitgeschätzten a-posteriori Standardabweichungen  $\hat{s}_A$  und  $\hat{s}_D$  der ursprünglichen Beobachtungen, gemessen im Amplituden- und Entfernungskanal, sowie die nach der Ausgleichung im Zuge der Fehlerrechnung bestimmten a-posteriori Standardabweichungen  $\hat{s}_{\hat{A}}$  und  $\hat{s}_{\hat{D}}$  der ausgeglichenen Messwerte beider Beobachtungsgruppen sind in **Tabelle 11.7** zusammengefasst.

Für die Bildkoordinatenmessungen  $\mathbf{x}'_{ij}$  der Kugelmittelpunkte  $\mathbf{X}_j$  fällt die geschätzte a-posteriori Standardabweichung  $\hat{s}_A$  mit ca.  $1/35$  px etwas besser aus als die im **Abschnitt 11.2.2** bestimmten **LSM**-Standardabweichungen der Translationsparameter. Die Messung der Distanzen  $D_{ijk}(\mathbf{x}')$  zwischen den Projektionszentren  $\mathbf{X}_{0i}$  und den Punkten  $\mathbf{X}_{jk}$  der Kugeloberflächen liegen mit einer a-posteriori Standardabweichung  $\hat{s}_D$  von ca.  $9 \text{ mm}$  im bereits durch Voruntersuchungen bestimmten Genauigkeitsniveau einer wiederholten **3D**-Kamerastreckenmessung (**Abschnitt 11.6**; **M. Hempel 2007**). Die



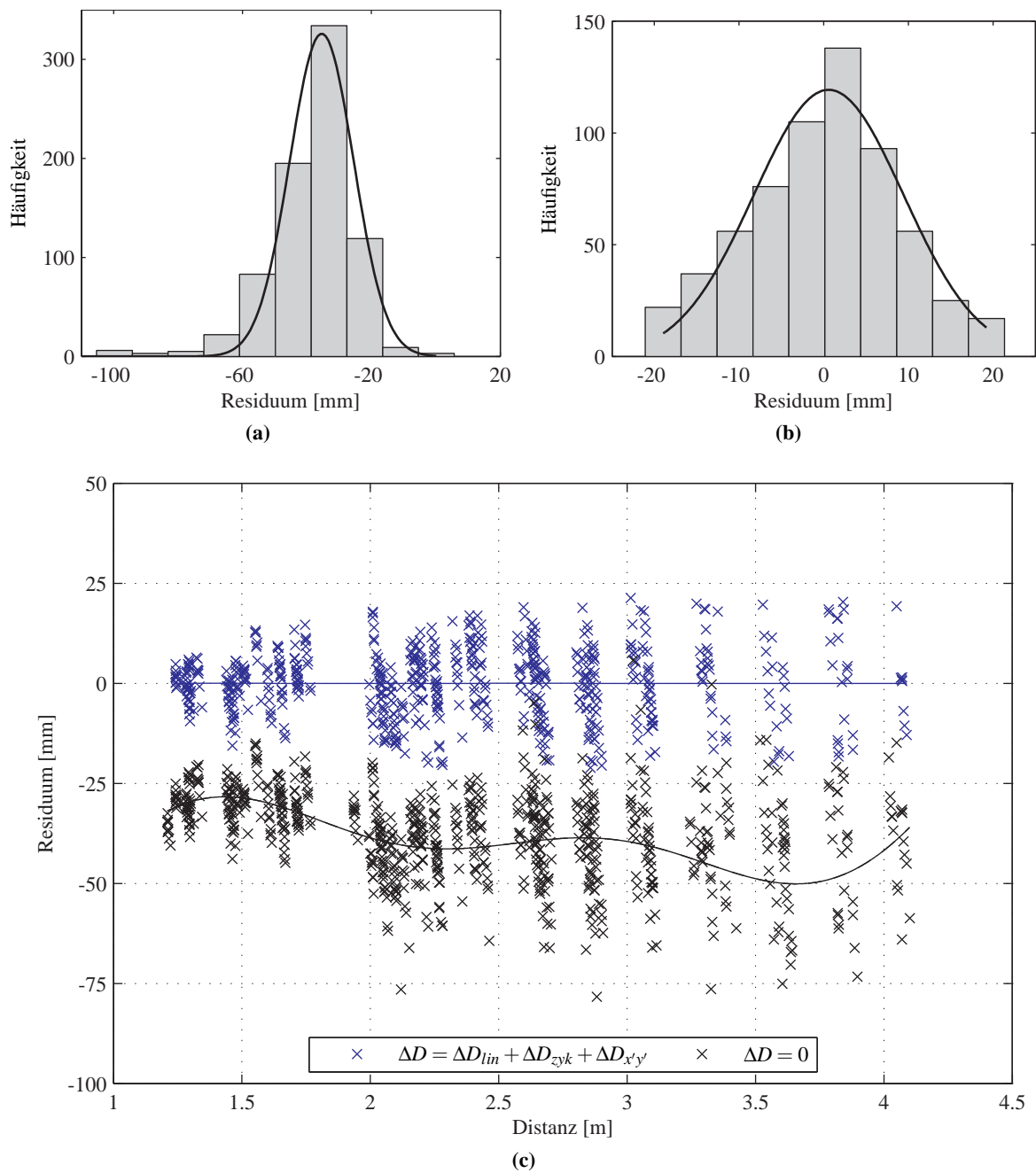
**Abb. 11.16:** Die verbleibenden Restklaffungen  $v_A$  der Bildkoordinatenmessungen  $\mathbf{x}'_{ij}$  sind ohne erkennbare systematische Effekte (a) mit einem quadratischen Mittel der a-posteriori Standardabweichung  $\hat{s}_{v_A}$  in beiden Bildkoordinatenrichtungen von ca.  $1/40$  px normalverteilt um den Erwartungswert  $\mu = -0,02 \mu\text{m}$  (b). Überhöhung: 200.

a-posteriori Standardabweichung  $\hat{s}_0$  der Gewichtseinheit liegt nahe der a-priori angesetzten Konstante  $s_0$ , in diesem Fall also nahe Eins. Entsprechend den Ausführungen des [Abschnitts 4.6](#) fällt die geschätzte Gewichtung weder zu optimistisch noch zu pessimistisch aus. Vielmehr konnten die Genauigkeitsverhältnisse der heterogenen Beobachtungen optimal bestimmt und anschließend im stochastischen Modell  $\Sigma_{II}$  berücksichtigt werden. Folglich entsprechen die in [Tabelle 11.7](#) gelisteten a-posteriori Standardabweichungen  $\hat{s}_{A,D}$  den a-priori Messgenauigkeiten  $s_{A,D}$  der ursprünglichen Bildkoordinaten- und Distanzbeobachtungen (Gleichung 4.31).

Die quadratischen Mittel der Standardabweichungen  $\hat{s}_{\hat{A}}$  und  $\hat{s}_{\hat{D}}$  der ausgeglichenen Beobachtungen, welche frei von systematischen Fehlern sind, liegen bei ca.  $1/90$  px für die Bildkoordinatenmessungen und bei ca. 1 mm für die Distanzmessungen. Diese inneren Genauigkeitswerte bestätigen zusammen mit den in [Abschnitt 11.8.7](#) analysierten verbleibenden zufälligen Messabweichungen, dass das vorliegende Beobachtungsmaterial adäquat durch das aufgestellte funktionale und stochastische Modell vermittelnd ausgeglichen wird.

**Tab. 11.7:** A-posteriori Standardabweichungen  $\hat{s}$  der ursprünglichen bzw. ausgeglichenen Beobachtungen sowie der Gewichtseinheit.

$\hat{s}_A$	$\hat{s}_{\hat{A}}$	$\hat{s}_D$	$\hat{s}_{\hat{D}}$	$\hat{s}_0$
1,16 $\mu\text{m}$	0,50 $\mu\text{m}$	8,713 mm	0,902 mm	1,00063



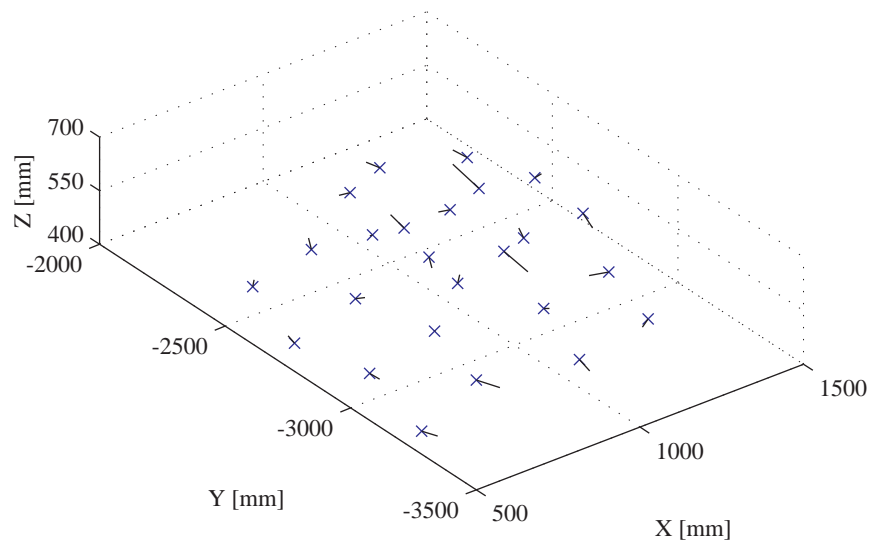
**Abb. 11.17:** Die Residuen der Distanzbeobachtungen weisen ohne eine Berücksichtigung von distanzabhängigen Fehlern deutlich sichtbare systematische Effekte auf (a und c). Die mittlere quadratische Abweichung  $RMS$  beläuft sich auf 40,86 mm. Werden die Distanzkorrekturterme in der integrierten Bündelblockausgleichung mitgeschätzt, streuen die verbleibenden Residuen ohne interpretierbare systematische Effekte mit  $\hat{\sigma}_{v_D} = 8,66$  mm normalverteilt um den Erwartungswert  $\mu = 0,58$  mm (b und c).

### 11.8.9 Äußere Genauigkeit

In den vorangegangenen Abschnitten wurden die a-posteriori Standardabweichungen der Beobachtungen und Unbekannten angegeben, um die Qualität der Ausgleichung zu beurteilen. Diese innere Genauigkeit resultiert aus den Abweichungen der einzelnen unter konstant gehaltenen Versuchsbedingungen gemessenen Beobachtungen zu ihrem Mittelwert (Abschnitt 4.5.1). Sie spiegelt folglich nicht die tatsächliche physikalische Messgenauigkeit wider. Zur Spezifikation der äußeren Genauigkeit ist vielmehr ein Vergleich mit Referenzdaten übergeordneter Genauigkeit notwendig.

Im Rahmen dieser Arbeit werden die Ergebnisse einer photogrammetrischen Bündelblockausgleichung auf der Basis von in den Bildern einer hochauflösenden digitalen Spiegelreflexkamera Nikon D300 mittels Ellipsenoperator gemessenen Bildkoordinaten zur Analyse herangezogen. Eine ausgleichende Helmerttransformation bildet die beiden Koordinatensysteme aufeinander ab und ermöglicht den direkten Vergleich der Neupunktkoordinaten. Die daraus resultierenden Differenzvektoren sind in [Abbildung 11.18](#) dargestellt und weisen keine verbleibenden Restsystematiken auf. Die mittlere quadratische Abweichung  $RMS_{\Delta X}$  aller Koordinatendifferenzen beträgt 5,0 mm, die minimale Abweichung 0,7 mm, die maximale Abweichung 10,4 mm.

Die ermittelte Objektpunktkoordinatengenauigkeit fällt mit ca. einem halben Zentimeter im Vergleich zu den von [Böhm und Pattinson \(2010\)](#) sowie [Lichti u. a. \(2010\)](#) publizierten Ergebnissen sehr gut aus. Es wird deutlich, dass 3D-Kameras nach der Korrektur der teils durch starke systematische Effekte beeinflussten Messdaten genaue und zuverlässige Ergebnisse liefern können. Ein Vergleich mit den deutlich kleineren quadratischen Mittelwerten der a-posteriori Standardabweichungen der Neupunktkoordinaten zur Beurteilung der inneren Genauigkeit belegt jedoch auch die Unvollständigkeit des aufgestellten Distanzkorrekturmodells ([Abschnitt 11.8.6](#)).



**Abb. 11.18:** Vergleich der geschätzten 3D-Kugelmittelpunktkoordinaten mit einem photogrammetrisch bestimmten Referenzdatensatz. Die mittlere quadratische Abweichung  $RMS_{\Delta X}$  beträgt 5,02 mm  
Überhöhung: 10.

### 11.8.10 Verfahrensvergleich

Die simultane Ausgleichung aller zur Verfügung stehenden 3D-Kamerainformationen in einem gemeinsamen integrierten Ansatz stellt eine Forderung an die entwickelte Bündelblockausgleichung dar. Der folgende Abschnitt untersucht, inwieweit eine integrierte Kalibrieroutine Vorteile gegenüber einem quasi-simultanen Ansatz bringt. Hierfür werden die im Amplitudenkanal gemessenen Bildkoordinaten in einer mit lediglich den inneren und äußeren Orientierungsparametern sowie den Mittelpunktkoordinaten parametrisierten Bündelblockausgleichung ausgewertet<sup>6</sup>. Auf der Basis der geschätzten Unbekannten werden anschließend die Distanzkorrekturterme ausgleichend unter Verwendung des reduzierten Distanzbeobachtungsmaterials ermittelt. Die Eignung der aufgestellten funktionalen und stochastischen Zusammenhänge kann nun über die Analyse der Residuen, der a-posteriori Standardabweichungen sowie der Differenzen der Kugelmittelpunkte zu Referenzkoordinaten beurteilt werden.

Die Unterschiede in den Beträgen sowie den Genauigkeiten der Orientierungsparameter selbst sind gering. Die Werte der Distanzkalibrierparameter variieren zwar im Vergleich zueinander, die Größenordnung ihrer Summe  $\Delta D$  und folglich auch die Auswirkungen des als gesamt zu betrachtenden Distanzkalibriermodells sind jedoch für alle Entfernungsbereiche ähnlich (vgl. [Abbildung 11.15](#)). Auch an dieser Stelle muss für ausführlichere Untersuchungen auf zukünftige Arbeiten verwiesen werden.

Die [Tabelle 11.8](#) fasst die in den [Abschnitten 11.8.7 bis 11.8.9](#) ausführlich besprochenen Qualitätskriterien der integrierten Bündelblockausgleichung zusammen und stellt sie den aus einem sequentiellen Ansatz produzierten Ergebnissen vergleichend gegenüber. Die a-posteriori Standardabweichungen  $\hat{s}_{v_{A,D}}$  der verbleibenden Residuen  $v_A$  der Bildkoordinatenmessungen und der Residuen  $v_D$  der Distanzmessungen verbessern sich für den integrierten Fall um jeweils ca. 15 %; die Erwartungswerte  $\mu_A$  und  $\mu_D$  der Standardnormalverteilungen tendieren deutlich stärker gegen Null. Auch die a-posteriori Standardabweichungen  $\hat{s}_{\hat{A},\hat{D}}$  der ausgeglichenen Beobachtungen bestätigen dies durch eine leichte Verbesserung in den Bildkoordinatenmessungen sowie eine Verdoppelung des Genauigkeitsniveaus für die Distanzmessungen. Schließlich zeigen die quadratischen Mittelwerte  $RMS_{\Delta X}$  der Koordinatendifferenzen der Kugelmittelpunkte eine marginale aber dennoch erkennbare Verbesserung der äußeren Genauigkeit einer simultanen photogrammetrischen und Distanzkalibrierung im Vergleich zu einer auf reinen Bildpunktmessungen basierenden photogrammetrischen Auswertung. Alle Sachverhalte sprechen für eine bessere Anlehnung des aufgestellten Modells an die geometrisch-physikalische Wirklichkeit und zeigen den Nutzen einer gemeinsamen Ausgleichung durch die implizite Berücksichtigung der Wechselwirkungen aller Modellparameter in einem integrierten Ansatz.

<sup>6</sup> Die Steuerparameter (z. B. Abbruchkriterien, Signifikanzniveaus usw.) der sequentiell durchgeführten Ausgleichungen des Amplituden- und Distanzkanals sind identisch mit denen der integriert gerechneten Bündelblockausgleichung.

**Tab. 11.8:** Vergleich der a-posteriori Standardabweichungen  $\hat{s}$  der Residuen, der Gewichtseinheit, der ursprünglichen und ausgeglichenen Beobachtungen sowie der RMS-Abweichungen der Koordinatendifferenzen zur Beurteilung der äußeren Genauigkeit für den Fall einer 2-Stufen-Kalibrierung (Zeile 1 und 2) und einer simultanen Kalibrierung (Zeile 3). Alle Angaben in [mm].

Beobachtung	$\hat{s}_{v_A}$ $\mu_A$	$\hat{s}_{v_D}$ $\mu_D$	$\hat{s}_0$	$\hat{s}_A$	$\hat{s}_D$	$\hat{s}_{\hat{A}}$	$\hat{s}_{\hat{D}}$	$RMS_{\Delta x}$
$\mathbf{x}'_{ij}$	$1,34e^{-3}$ $-0,07e^{-3}$	-	$9,93e^{-4}$	1,00	-	$5,37e^{-4}$	-	5,12
$D_{ijk}$	-	10,24 0,83	12,72	-	1,00	-	1,62	-
$\mathbf{x}'_{ij}$ und $D_{ijk}$	$1,15e^{-3}$ $-0,02e^{-3}$	8,66 0,58	1,00	$1,16e^{-3}$	8,71	$5,02e^{-4}$	0,90	4,95

## 11.9 Diskussion

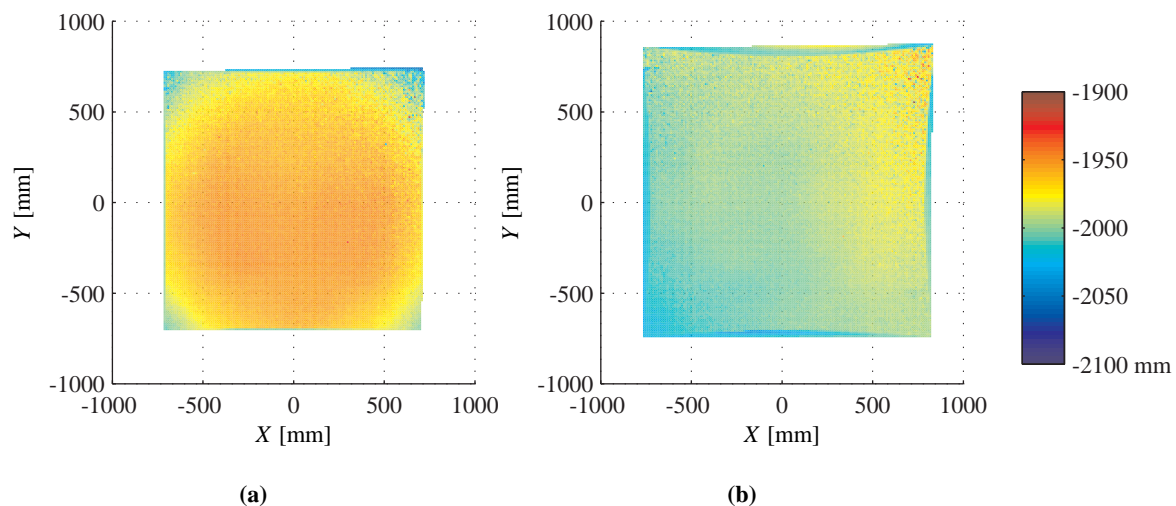
Die vorgestellte Bündelblockausgleichung ermöglicht die Bestimmung der exakten Kamerageometrie sowie die Schätzung der distanzmessspezifischen Korrekturparameter. Der entwickelte Ansatz setzt die Hauptforderung einer Integration aller zu Verfügung stehenden Messwerte in einem gemeinsamen funktional-stochastischen Zusammenhang über die Formulierung als eine Ausgleichung nach vermittelnden Beobachtungen mit Bedingungen zwischen den Unbekannten und automatischer Schätzung der Varianzkomponenten vollständig um. Basierend auf dem Prinzip einer Selbstkalibrierung werden neben den hauptsächlich interessierenden Parametern der inneren Orientierung und der Distanzkorrektur die unbekannte und zeitlich stabile Geometrie eines kompakt entworfenen 3D-Referenzfelds bestimmt. Die zur Distanzkalibrierung notwendige und instrumentell recht aufwendige Ermittlung von Referenzstrecken übergeordneter Genauigkeit kann somit entfallen.

Die Analyse der verbleibenden und normalverteilten Residuen der Beobachtungen sowie aller zur Verfügung stehenden a-posteriori Standardabweichungen bestätigen zum einen die Richtigkeit des aufgestellten funktionalen Zusammenhangs, vor allem im Bezug auf die integrierte geometrische Formulierung beider Beobachtungsarten. Sie offenbart zum anderen allerdings auch noch vorhandene Schwächen des Distanzkorrekturmodells. Da aufgrund der recht jungen Tiefenbildtechnologie noch nicht alle distanzmessspezifischen Fehler wie bspw. die *Scattering*-Effekte modelliert werden konnten, ist die Möglichkeit der einfachen Erweiterung des funktionalen Modells zur auch zukünftigen Nutzung des entwickelten Verfahrens notwendig. Die Formulierung als ein klassisches Kleinste-Quadrate-Problem und die anschließende modulare Programmierung setzen diese Forderung konsequent um.

Obwohl das im Rahmen dieser Arbeit aufgestellte Distanzkorrekturmodell nicht vollständig ist, konnte vor allem durch die Bestimmung der linearen, zyklischen und signalwegeffektbedingten Fehleranteile einer 3D-Kamerastreckenmessung die Qualität der 3D-Neupunktkoordinaten deutlich ver-

bessert werden. Deren äußere Genauigkeiten betragen nach dem Anbringen der ermittelten Korrekturparameter 5 mm.

Das folgende Beispiel zeigt die Auswirkungen einer 3D-Kamerakalibrierung anhand der Aufnahme einer Wand senkrecht zur Aufnahmerichtung in einem Abstand von etwa zwei Metern. In [Abbildung 11.19a](#) sind die gerechneten  $XY$ -Objektpunktkoordinaten mit ihrer farbkodierten  $Z$ -Koordinate dargestellt. Deutlich erkennbar ist die für eine unkalibrierte 3D-Kamera typisch konvexe Form einer eigentlich ebenen Wand. Unter der Berücksichtigung der Orientierungs- und Distanzkalibrierungsparameter verschwinden diese systematischen Effekte nahezu gänzlich. Einzig Fehlpixel in der oberen linken Sensorecke sowie leicht systematische Effekte an der rechten Seite der Wand schmälern den visuellen Eindruck.



**Abb. 11.19:** Veranschaulichung der Resultate der Simultankalibrierung anhand der Aufnahme einer weißen ebenen Wand in einem Abstand von zwei Metern senkrecht zur Aufnahmerichtung. (a) zeigt die  $XY$ -Objektpunktkoordinaten ohne und (b) unter Berücksichtigung der Orientierungs- und Distanzkalibrierungsparameter als Funktion ihrer farbkodierten  $Z$ -Koordinate.

Die während der Entwicklung aufgetretenen Fragestellungen können Untersuchungsgegenstände für zukünftige Arbeiten sein und werden kurz zusammengefasst:

- Die Anzahl der unbekannt und für die Kalibrierung prinzipiell nicht notwendigen 3D-Koordinaten der Kugeloberflächenpunkte ist sehr groß ([Abschnitt 11.3.5](#)). Für die Stabilität des Gleichungssystems ist eine Formulierung des funktionalen Zusammenhangs ohne die diskreten Objektpunktkoordinaten wünschenswert. In der zur automatischen Näherungsberechnung aufgestellten Quasi-Beobachtungsgleichung (11.7) wurde dieser Zusammenhang über den Schnitt eines Projektionsstrahls mit der Kugeloberfläche bereits analytisch hergestellt. Durch eine Umstellung nach  $(D_{ijk} + \Delta D_{ijk})$  kann die quadratische Gleichung (11.7) zu einer direkten Beobachtungsgleichung umgeformt und in die Bündelblockausgleichung integriert

werden. Lediglich die damit einhergehenden Mehrdeutigkeiten sind im Ausgleichungsmodell zu berücksichtigen.

- Die Integration adaptiver Varianzkomponenten zur Berücksichtigung eines konstanten und eines linear von der Distanzmessung selbst abhängigen Fehleranteils kann vorgenommen werden ([Abschnitt 11.6](#)).
- Eine vergleichende Analyse der Kalibrierergebnisse baugleicher 3D-Kameras bzw. Geräte unterschiedlicher Hersteller ist denkbar ([Abschnitt 11.8.1](#)).
- Die Erweiterung der Experimentalkonfiguration um 3D-Kameradaten unterschiedlicher Modulationsfrequenzen und Integrationszeiten ermöglicht die Bestimmung anwendungsspezifischer Kalibrierparametersätze ebenso wie die Erweiterung um Aufnahmen über den gesamten Messbereich unter der Verwendung eines entsprechend dimensionierten Referenzfelds ([Abschnitt 11.8.1](#)).
- Prinzipbedingt wird das Beobachtungsmaterial teils ausgedünnt. Die Auswirkungen dieser Reduktion auf die Genauigkeit und Zuverlässigkeit der Ausgleichung sind detailliert zu untersuchen ([Abschnitt 11.8.2](#)).
- Die zeitliche Stabilität der Kalibrierparameter ist von Interesse ([Abschnitt 11.8.5](#)).



# IV

Bewegung



## Verfahren zur Erfassung menschlicher Bewegungen

Die Erfassung von Objektbewegungen ist in zahlreichen Anwendungen und unterschiedlichen Disziplinen von großer Bedeutung. Im [Kapitel 7](#) sind bereits mögliche Szenarien kurz skizziert worden, welche von technisch-industriellen Anwendungsfeldern über Einsatzgebiete in Medien- und Unterhaltungsbranchen, im Umweltmonitoring oder bei Überwachungs- und Sicherheitsaufgaben bis hin zu medizinischen Einsatzgebieten reichen. Die vorliegende Arbeit fokussiert speziell auf die Analyse menschlicher Bewegungen. Als praktisches Anwendungsbeispiel mit großer Relevanz in Psychologie und Erziehungswissenschaften dient zur Dokumentation des Potentials der entwickelten Verfahren die Erfassung von Bewegungsabläufen in der Analyse nonverbaler Kommunikation in Interaktionsverläufen von Personen, welches später im [Kapitel 15](#) konkretisiert wird.

Das menschliche Bewegungsverhalten wird üblicherweise über Feldbeobachtungen vor Ort erfasst, die relevanten Größen simultan mit sogenannten Papier- und Bleistift-Verfahren dokumentiert und anschließend manuell ausgewertet (z. B. [LaFrance & Broadbent 1976](#); [Stoffert 1985](#); [York-Barr & Duke 2004](#)). Nicht nur in den Sozial- und Verhaltenswissenschaften wurde früh Kritik an dieser Vorgehensweise laut, da ein *teilnehmender Beobachter* stets indirekt Einfluss auf das Verhalten des Probanden nimmt ([Webb u. a. 1975](#)). Seit dem Aufkommen von für den Anwender preislich erschwinglicher analoger und später auch digitaler Videotechnik gegen Ende des letzten Jahrhunderts werden unterstützend Foto- und Videotechniken eingesetzt. Die Auswertung von Videosequenzen stellt in heutiger Zeit eines der leistungsfähigsten und vielseitigsten Werkzeuge zur Datenakquisition in vielen Bereichen der menschlichen Bewegungs- und Verhaltensanalysen dar. Videosequenzen können interaktiv oder automatisch bzw. automatisiert ausgewertet werden. Eine interaktive Auswertung nutzt zwar das Kontextverständnis des Operators, ist aber – neben dem erheblichen Aufwand der interaktiven Auswertung – auch mit Subjektivität, Generalisierung und Segmentbildung auf der Zeitskala verbunden. Eine weitgehende Automatisierung scheint daher gerade in datenintensiven Studien unter den Aspekten von Effizienz und Objektivität wünschenswert. Zur quantitativen Analyse können dabei Verfahren der *Computer Vision* und der Photogrammetrie einen wertvollen Beitrag leisten.

Die Entwicklung von Verfahren zur automatischen bildbasierten Rekonstruktion des menschlichen Körpers und dessen Bewegung spielt bereits seit Anfang der 1990er Jahre eine wichtige Rolle. Allein in den Jahren 2000 bis 2006 gab es auf diesem Gebiet über 350 wissenschaftliche Publikationen ([Moeslund u. a. 2006](#)). Die Ansätze reichen von der Erfassung flächenhafter Bildstrukturen, Silhouetten oder diskreten Merkmalen bis zu auf komplexen Mensch- und Bewegungsmodellen basierenden Methoden. Selbstverständlich existieren neben bildbasierten Verfahren zur Bewegungserfassung (*Motion Capturing*) und Bewegungsanalyse (*Motion Analysis*) auch zahlreiche andere Systemkonfi-

gurationen, wie bspw. Kraftmessplatten, Beschleunigungssensoren, Druckverteilungsmessungen und viele mehr. Die Dissertationsschrift von [Breiner \(2005\)](#) bietet eine detaillierte Übersicht verschiedenster Verfahren der Bewegungserfassung.

Das folgende Kapitel zeigt Verfahren zur Bewegungserfassung und gibt zusätzlich Beispiele für in der Literatur implementierte Methoden zur Extraktion von Bewegungsabläufen aus Bildsequenzen. Der [Abschnitt 12.1](#) behandelt zunächst kurz die Vorgehensweise bei einer interaktiven Auswertung von Videografiedaten. Verschiedene und entsprechend ihrer Instrumentalkonfiguration kategorisierte photogrammetrische Verfahren zur automatischen Bildsequenzanalyse mit dem Ziel der Extraktion menschlicher 2D- und 3D-Bewegungsinformationen werden im [Abschnitt 12.2](#) vorgestellt. Abschließend zeigt der [Abschnitt 12.3](#) kommerziell erhältliche Softwarepakete und Komplettlösungen zur Punktverfolgung.

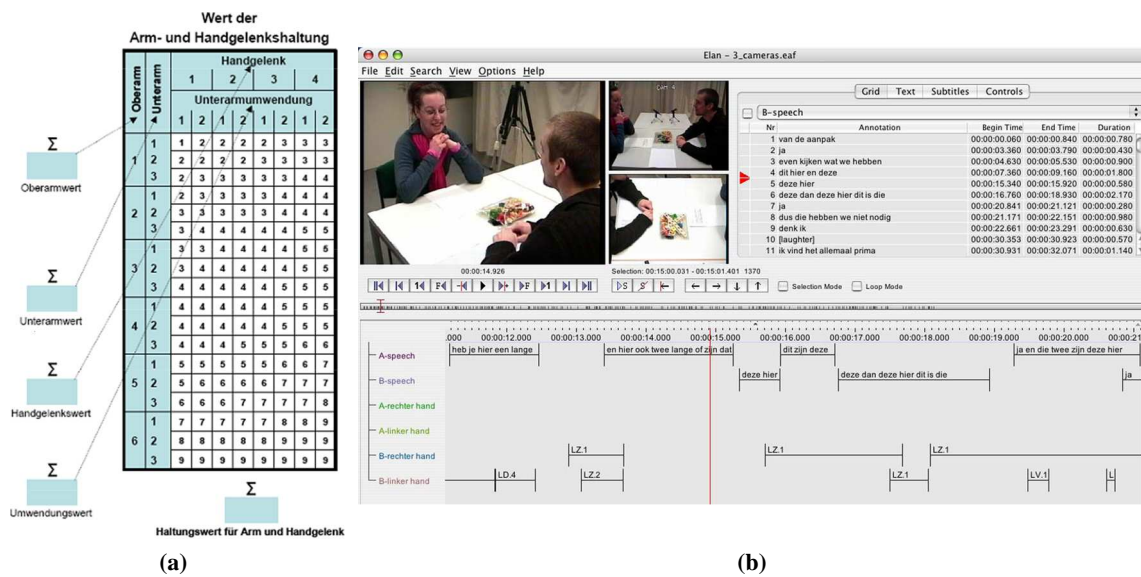
Für detailliertere Kategorisierungen und weiterführende Informationen sei zusätzlich auf ([Frischholz 1999](#); [Remondino 2006](#)) verwiesen.

## 12.1 Bewegungsinformation durch interaktive Videografie

Analoge oder digitale Videosequenzen menschlicher Bewegungen können an entsprechend ausgestatteten Arbeitsstationen interaktiv ausgewertet werden. Hierzu stehen Playbackstationen und Kodierungswerkzeuge zur Verfügung, die es erlauben, Bewegungsdaten zu sammeln und zu analysieren. [Grynszpan u. a. \(2003\)](#) erfassen auf diese Weise Parameter wie Gestik, Sprache, Gesichtsausdruck und Blick von autistischen Personen. [Kommer und Biermann \(2005\)](#) analysieren auf der Basis umfangreichen Videomaterials Gestik, Mimik und Umfeld von Probanden in Computerkursen. [Schmitter \(2005\)](#) stuft ungünstige Körperhaltungen am Arbeitsplatz für Körperbereiche wie Kopf, Extremitäten usw. rechnerunterstützt unter der Verwendung entsprechender Eingabemasken anhand standardisierter Körperbewegungsbereiche grob ein ([Abbildung 12.1a](#)).

Das Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung (Sankt Augustin, D) schlägt zur Erfassung und Bewertung physischer Arbeitsbelastungen eine unterstützende Videodokumentation vor ([BGIA 2006](#)). [Lausberg u. a. \(2007\)](#) und [Lausberg und Sloetjes \(2009\)](#) verwenden routinemäßig auf Videodaten angewendete Annotationssoftware (z. B. *ELAN*; [Abbildung 12.1b](#)) zur manuellen Detektion von Bewegungsaktivitäten bei Gesprächen zwischen bspw. einem Therapeuten und seinem Patient sowie zur anschließenden Interpretation der aufgetretenen Gestentypen (vgl. auch [Abschnitt 15.2](#)). [Doherty-Sneddon und Phelps \(2007\)](#) untersuchen ebenfalls über Videografieaufzeichnungen und nachträgliche manuelle Kodierung Effekte nonverbalen Verhaltens von Lehrkräften zu Schülern.

Der entscheidende Vorteil einer vollständig manuellen Analyse ist, dass die Interpretation der aufgenommenen Bewegungen im Sinne der jeweiligen Aufgabenstellung zumeist zuverlässig durch einen erfahrenen und routinierten Auswerter durchgeführt wird. Dem gegenüber steht jedoch ein enormer Zeitaufwand bei der vorherigen Erfassung von bspw. der Häufigkeit und des Zeitanteils eines Vorkommnisses.



**Abb. 12.1:** Interaktive Erfassung und manuelle Interpretation von Bewegungsabläufen: (a) Die Erhebungen von Arbeitsplatzbeobachtungen und die Einstufung ungünstiger Körperhaltungen erfolgt in Erfassungsbögen, teils rechner- und videogestützt (Schmitter 2005). (b) Screenshot der Annotationssoftware ELAN des Max-Planck-Institut für Psycholinguistik (Nijmegen, NL). Mehrere synchronisierte Videosequenzen aus unterschiedlichen Perspektiven (oben links) stehen für die durch einen Operator vorzunehmende Kodierung auf der Zeitachse (unten) zur Verfügung (Wittenburg u. a. 2006).

## 12.2 Bewegungsinformation durch automatische Bildsequenzanalyse

Es ist wünschenswert, zumindest Teilaspekte der Auswertung zu automatisieren, die bei einer interaktiven Analyse mit sehr hohem Zeitaufwand verbunden sind. Im einfachsten Fall wäre es eine reine Bewegungsdetektion diskreter Körperpunkte. In der *Computer Vision* und der Photogrammetrie werden zur Extraktion von Bewegungsinformationen aus Bildsequenzen häufig flächenbasierte Bildzuordnungsverfahren wie die Kreuzkorrelation, der Optische Fluss oder das *Least Squares Matching (LSM)* angewendet (Abschnitte 7.4 bis 7.6). Sie liefern den Verschiebungsvektor eines einzelnen Bildpunkts zwischen zwei betrachteten Zeitpunkten. Die Länge des geschätzten Bewegungsvektors ist sodann ein Kriterium, ob der betrachtete Körperpunkt zum entsprechenden Zeitpunkt eine relevante Bewegung erfuhr oder nicht. Findet dieser Zuordnungsprozess nicht nur für einen Zeitpunkt statt, sondern zwischen allen konsekutiven Bildern einer Bildsequenz, entstehen 2D-Trajektorien der Körperpunkte. In Abhängigkeit von der räumlichen und zeitlichen Auflösung der Kamera können so über automatische Verfahren der Bildverfolgung für beliebig viele Körperpunkte Bewegungsfelder bestimmt, die Trajektorien einzeln interpretiert oder durch bspw. die Segmentierung bewegter Bereiche weiter verarbeitet werden.

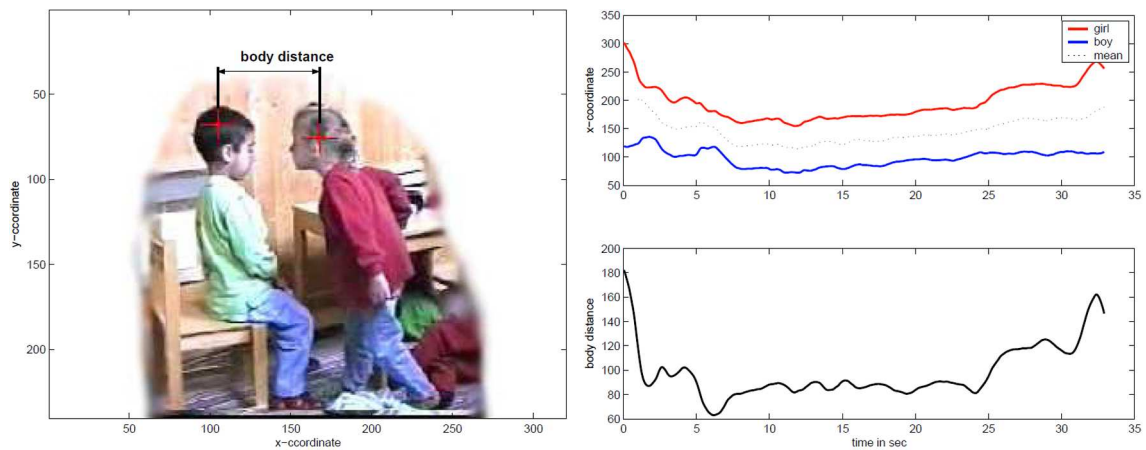
### 12.2.1 Monokulare Bildsequenzanalyse

Eine Bewegungsanalyse aus Einzelbildern reduziert den instrumentellen Aufwand erheblich, da lediglich eine Kamera zur Datenakquisition Verwendung findet, welche allenfalls bezüglich ihrer Aufnahmegeometrie einmalig zu kalibrieren ist ([Abschnitt 5.1.3](#)). Es erfordert aber zusätzliches Objektwissen über den zu erwartenden Bewegungsverlauf und/oder die Geometrie der aufgenommenen statischen und bewegten Objekte, um die in der monokularen Bildsequenz durchgeführten Messungen in den Objektraum übertragen zu können ([Abschnitt 5.3](#)).

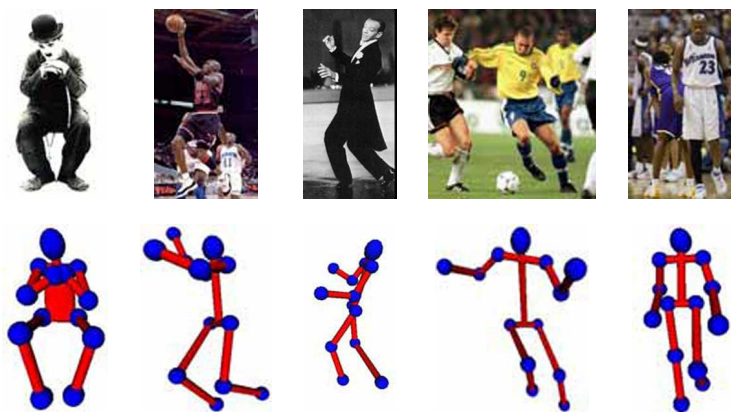
[Altmann u. a. \(2007\)](#) werten in einer Auflösung von  $(352 \times 288)$  px digitalisiertes VHS-Video material aus, um Aussagen über die Konfliktregulierung bei befreundeten Kindespaaren treffen zu können ([Abbildung 12.2](#)). In dem vorhandenen Datenmaterial werden u. a. die Köpfe mit 2D-LSM verfolgt ([Abschnitt 7.6.1](#)), die resultierenden Trajektorien analysiert und zur Ableitung der über die Zeit variierenden Körperdistanzen herangezogen (vgl. auch [Abschnitt 15.1](#)). Bei dieser Anwendung ist eine zweidimensionale Einzelbildanalyse möglich, da sich die Probanden bei der auszuwertenden Szene in annähernd einer Ebene bewegen. Weiterhin untersuchen die Interaktionsforscher vor allem die Synchronisation der Bewegungsabläufe. Eine Auswertung im Bildraum erscheint somit sinnvoll, da keine absoluten Maße von Interesse sind.

Den Übergang zur dreidimensionalen Rekonstruktion von Bewegungsabläufen aus Einzelbildsequenzen vollzieht [Remondino \(2004\)](#) bei der Auswertung von Basketballspielen aus Amateurvideos ([Abbildung 12.2](#)). Zwangsbedingungen für eine 3D-Rekonstruktion auf niedrigem Genauigkeitsniveau ergeben sich hier u. a. aus der Geometrie des Spielfelds und der Kenntnis der Längenverhältnisse von Gliedmaßen. Die Nutzbarkeit von Amateurvideos eröffnet dabei auch interessante Aspekte zur Auswertbarkeit historischer Aufnahmen ([Remondino 2003](#)). In ähnlicher Weise rekonstruieren [Howe u. a. \(1999\)](#) dreidimensionale menschliche Bewegungen aus 2D-Bildsequenzen. Ihr präsentiertes System basiert dabei auf aus Trainingsdatensätzen erlangten Vorkenntnissen der zu erwartenden Bewegungen. In einem probabilistischen Ansatz verwenden [Sidenbladh u. a. \(2000\)](#) ein generatives Transformationsmodell unter der Berücksichtigung von Gelenkentfernungen und Gelenkstellungen zur 3D-Verfolgung von Menschen in 2D-Videodaten. In ([Urtasun & Fua 2004](#)) werden menschliche Körper in 2D-Bildsequenzen auf der Basis zeitlicher Bewegungsmodelle verfolgt. [Dimitrijevic u. a. \(2005\)](#) detektieren und verfolgen menschliche Bewegungen in monokularen Bildsequenzen einer bewegten Kamera mittels spatiotemporaler Silhouettenzuordnung ([Abbildung 12.4](#)).

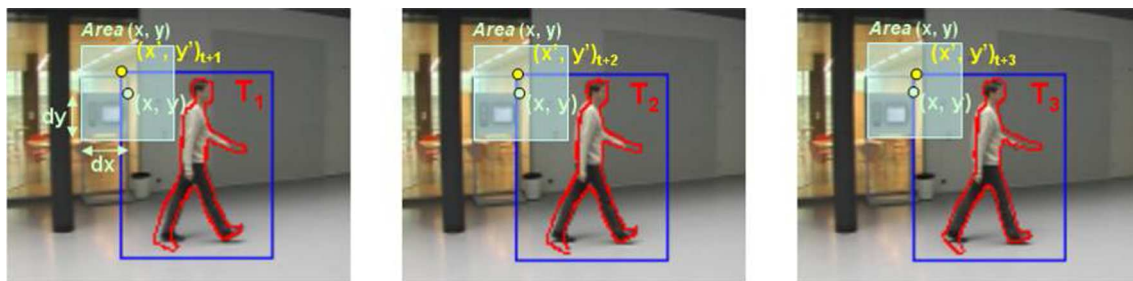
Die Entwicklung und Implementierung von Verfahren zur monokularen Bildsequenzanalyse ermöglicht einerseits die automatisierte Auswertung von bereits durch den Praktiker generierten Videografiedaten, andererseits eröffnet die Verwendung von lediglich einer Kamera eine einfache Möglichkeit der Datenakquisition für neu zu entwerfende Experimentalkonfigurationen. Aufgrund der damit einhergehenden Einschränkungen bei der Übertragbarkeit der gemessenen 2D-Daten in den 3D-Objektraum ist diese Vorgehensweise allerdings nur für spezielle Anwendungen unter der Einbeziehung gewisser räumlicher und zeitlicher Modellannahmen geeignet.



**Abb. 12.2:** 2D-Punktverfolgung in Einzelbildsequenzen: Die interpersonale Distanz zweier sich annähernd in einer Ebene bewogender Kinder kann über die 2D-Verfolgung der Köpfe approximiert werden (Altmann u. a. 2007).



**Abb. 12.3:** Beispiele für die Rekonstruktion menschlicher Körper(-bewegungen) aus Einzelbildern bzw. Videobildsequenzen alter Fernsehaufnahmen: In der oberen Reihe sind die Eingabedaten, in der unteren Reihe die rekonstruierten und durch Zylinder und Ellipsoide verbundene Objektpunkte dargestellt (Remondino 2006).



**Abb. 12.4:** Spatiotemporale Silhouettenbildzuordnung in konsekutiven Bildern einer monokularen Bildsequenz (Dimitrijevic u. a. 2005).

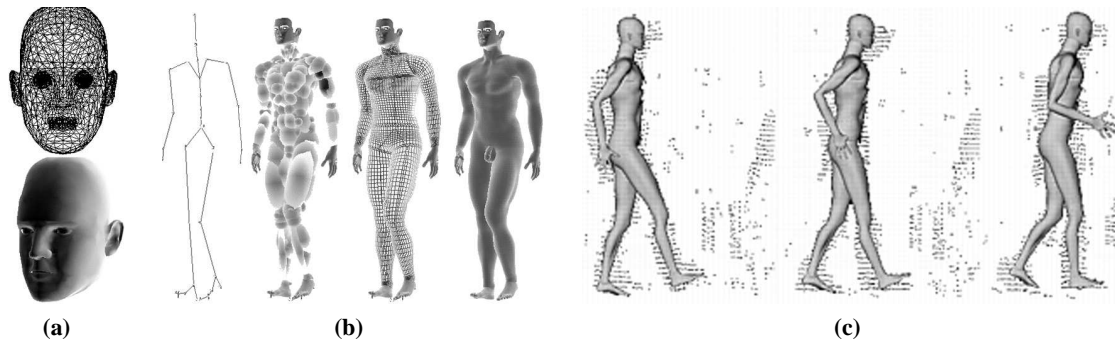


### 12.2.2 Multiokulare Bildsequenzanalyse

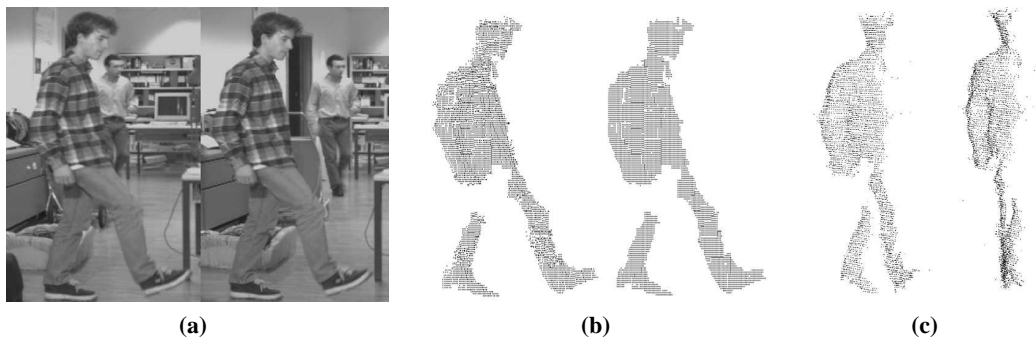
Eine naheliegende Erweiterung der zunächst auf eine zweidimensionale Projektion der Bewegungsinformation beschränkten videografischen Bildsequenzanalyse zur 3D-Bewegungsanalyse ergibt sich durch die Verwendung stereoskopischer Aufnahmesysteme mit zwei möglichst konvergent ausgerichteten Kameras (Abschnitt 5.4). Sensortechnisch bedingt dies lediglich die Synchronisation und Kalibrierung der beiden Kameras. Bei der Auswertung ergibt sich zusätzlich zur Verfolgung natürlicher oder signalisierter Objektpunkte auch die Notwendigkeit der räumlichen Zuordnung korrespondierender Trajektorien, extrahiert aus zeitlich synchronisierten Bildsequenzen. Die Reihenfolge dieser beiden Verarbeitungsschritte kann der Zuordnung auf der Zeitachse (Netzsch & Jähne 1995) oder der stereoskopischen Zuordnung (Maas u. a. 1993) Priorität geben; Willneff und Grün (2002) zeigen hierfür einen integrierten Ansatz. Je nach Aufgabenstellung kann die Lösung der spatiotemporalen Zuordnung hohe Komplexität aufweisen und führt nicht immer zu fehlerfreien Lösungen. In vielen Fällen ist dann eine Konzeption multiokularer Messsysteme aus drei oder mehr Kameras notwendig (Maas 2005b). Echtzeitfähigkeit erreichen hier nur Systeme, die sich auf die Bestimmung der 3D-Trajektorien einer begrenzten Anzahl durch retroreflektierende Zielmarken signalisierter Punkte beschränken (Abschnitt 12.3).

In (Gavrila 1996) wird der Prototyp eines Mehrkamarasystems zur vollständigen Verfolgung von Ganzkörperbewegungen vorgestellt. Zu jedem Zeitpunkt werden die Parameter eines virtuellen Menschmodells in die Silhouette einer sich bewegenden Person, extrahiert aus den aus verschiedenen Blickrichtungen aufgenommenen Bildern, eingepasst. Das Resultat der Verarbeitungskette ist eine zeitlich aufgelöste, dreidimensionale Repräsentation menschlicher Bewegungsvorgänge. Fua u. a. (2000) passen ebenfalls generische Kopf- und Körpermodelle in 3D-Daten ein, um so den menschlichen Körper und dessen Bewegung zu rekonstruieren (Abbildung 12.5). Hier werden aber im Gegensatz zur berührungslosen Vorgehensweise von Gavrila die 3D-Punktwolken aus der Verfolgung von an repräsentativen Körperpunkten angebrachten und durch acht Infrarotkameras beobachteten retroreflektierenden Markern berechnet. D'Apuzzo (2003) zeigt eine Anwendung der Bewegungsanalyse auf der Basis eines Mehrkamarasystems zur dreidimensionalen Beschreibung der Bewegung einer gehenden Person für Zwecke der Computeranimation, die auf der Zuordnung vorhandener Objekttextur basiert (Abbildung 12.6). In Bildtripeln eines trinokularen Kamerasystems werden zunächst durch 2D-LSM Korrespondenzen für Punkte der Körperoberfläche gesucht, auf deren Basis im räumlichen Vorwärtsschnitt 3D-Koordinaten bestimmt werden (Abschnitte 5.2 und 7.6.1). In den Bildsequenzen der einzelnen Kameras werden – ebenfalls durch 2D-LSM – dieselben Punkte verfolgt. Anschließend erfolgt eine Konsistenzanalyse auf Basis der Redundanz in der spatiotemporalen Zuordnung. Ausreißer in den auf diese Weise erhaltenen Trajektorien werden über Schwellenwerte für Geschwindigkeit und Beschleunigung sowie Uniformitätsanalysen im Vektorfeld eliminiert. Vedula u. a. (2005) implementieren einen vergleichbaren Ansatz, um den mit 17 Kameras aufgenommenen Bewegungsablauf einer Tänzerin markerlos zu modellieren (Abbildung 12.7). Als Verfahren zur flächenbasierten Zuordnung texturierter Oberflächen kommt der Optische Fluss zum Einsatz (Abschnitt 7.5).

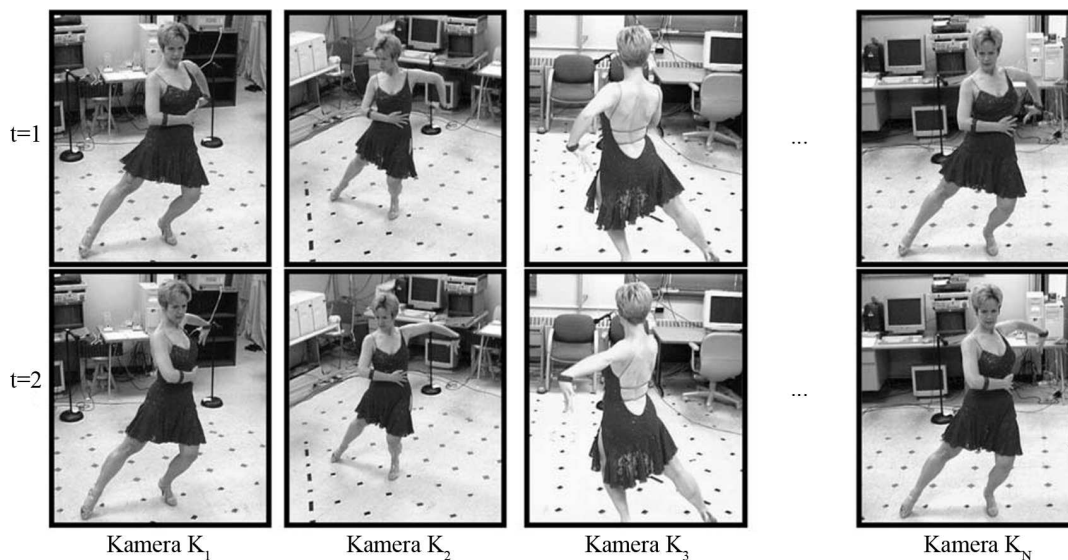




**Abb. 12.5:** Menschliche Körper- und Bewegungsrekonstruktion (Fua u. a. 2000): (a) Drahtgitter- und Oberflächenkopfmodelle. (b) Verschiedene Ganzkörpermodellrepräsentationen. (c) In aus der Bildzuordnung resultierende Punktwolke gefittetes 3D-Modell.



**Abb. 12.6:** Berührungslose 3D-Bewegungsanalyse auf der Basis eines Mehrkamerasystems zur Generierung eines 3D-Bewegungsvektorfelds einer gehenden Person für Zwecke der Computeranimation (D'Apuzzo 2003): (a) Bildpaar. (b) Zugeordnete Bildpunkte. (c) 3D-Punktwolke.



**Abb. 12.7:** Berührungslose 3D-Rekonstruktion einer Tänzerin aus konsekutiven Mehrkameraaufnahmen mittels Verfahren des Optischen Flusses (nach Vedula u. a. 2005).

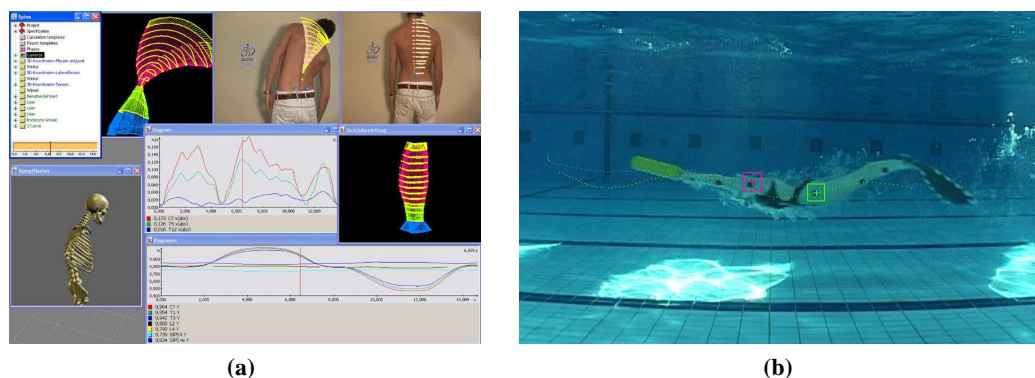
Multiokulare Systeme bilden die Basis für detaillierte Bewegungsanalysen und ermöglichen unter der Berücksichtigung der notwendigen Aufnahmekonfiguration ([Abschnitt 5.4](#)) und möglicher Korrespondenzprobleme ([Abschnitt 7.3](#)) die vollständige dreidimensionale Abbildung von Bewegungsvorgängen. Aktuell bilden sie so die Grundlage für kommerziell erhältliche Komplettsysteme zur Erfassung, Rekonstruktion und Analyse menschlicher Bewegungen.

## 12.3 Trackingsysteme

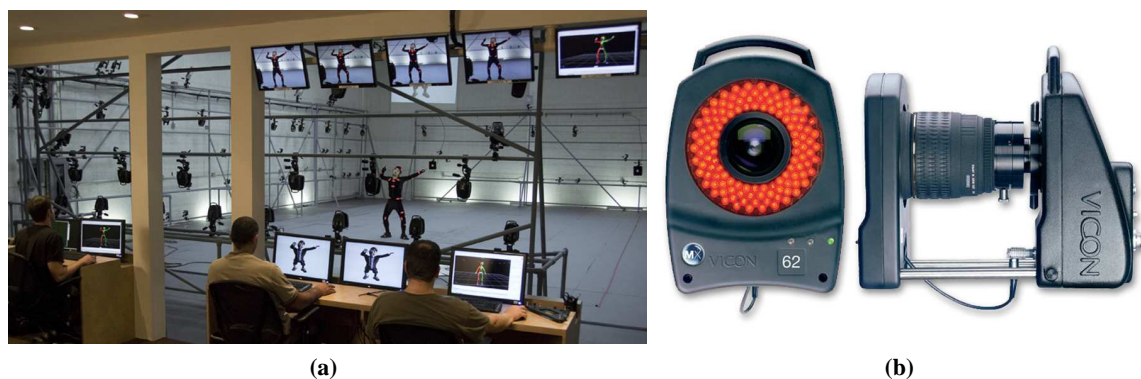
Optische Systeme zur vollständigen Erfassung von Bewegungsabläufen (*Motion Capturing Systems*) sind meist in speziellen Aufnahmestudios installierte Mehrkamarasysteme, die durch eine optimale Anordnung der Kameras rund um das aufzunehmende Objekt den gesamten Bewegungsbereich erfassen ([Abbildung 12.9](#)). Die Anzahl der verwendeten Kameras liegt hierbei in der Regel deutlich über zehn. Neben den digitalen Aufnahmegeräten und den entsprechenden Schnittstellen zur Synchronisation ist ein bewegungsanalytisches Messsystem zusätzlich mit auf einem Computer installierter Software zur Kameraansteuerung und Bilddatenauswertung ausgestattet. Solche in verschiedenen Ausführungen kommerziell vertriebene Softwarepakete realisieren die Detektion, Lokalisierung und Verfolgung von Objekten in Bildsequenzen. Beispielhaft seien an dieser Stelle die Programme *Cortex* der in diesem Bereich führenden US-amerikanischen Firma [Motion Analysis Corporation \(2011; \[Abbildung 12.10a\]\(#\)\)](#) und *Simi MatchiX* für markerlose Punktverfolgung der deutschen Firma [SIMI Reality Motion Systems GmbH \(2011\)](#) genannt.

Derzeitig kommerziell verfügbare Verfahren zur Bewegungserfassung arbeiten mit aktiven bzw. passiven Positionsmarkern, die an möglichst vielen repräsentativen Körperstellen wie z. B. Handrücken, Schultergelenke, Ellenbogen, Hüfte oder Knöchel angebracht sind. Sollen Mimik und Feingestik ebenfalls erfasst werden, sind weitere Marker im Gesicht bzw. an den Fingergliedern erforderlich. Solch eine Signalisierung mittels bspw. farbigen Leuchtdioden, im [nahen Infrarot \(NIR\)](#) retroreflektierenden Zielmarken, konzentrischen Kreisen oder kodierten Zielmarken vereinfacht die Detektion und Identifikation der Marker im Bild ([Maas 1997](#)) und gewährleistet ein zuverlässiges, teils echtzeitfähiges Messsystem ([Abbildung 12.10a](#)).

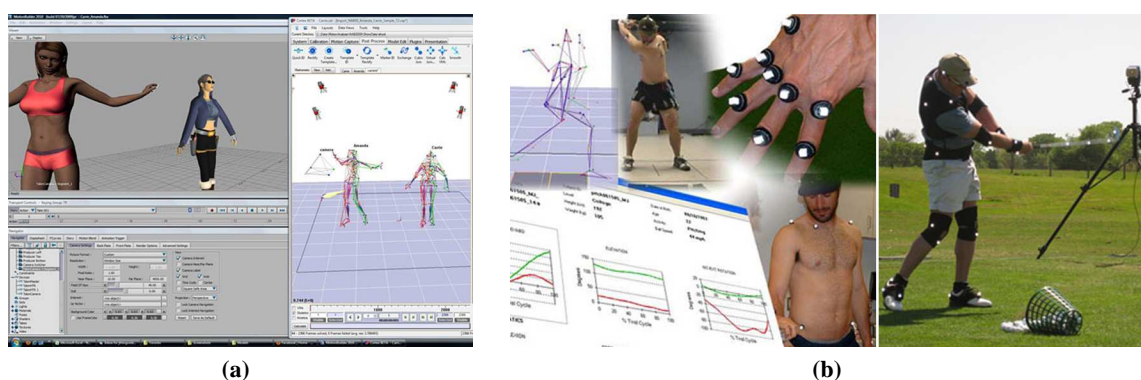
In der Praxis sind die Zielmarken bereits an speziellen Kleidungsstücken oder Ganzkörperanzügen angenäht, um das zeitaufwendige Anbringen einzelner Positionsmarker zu vermeiden ([Abbildungen 12.9a](#) und [12.10b](#); vgl. auch [Breiner 2005](#)). Für verschiedene Anwendungen ist eine solche Signalisierung jedoch nicht möglich bzw. nicht gewünscht. In der Automobilindustrie werden klebende Zielmarken auf neuen Produkten nicht toleriert, im Sport verhindern spezielle Kleidungsstücke die Anwendung unter Wettkampfbedingungen, in der Verhaltensforschung beeinflussen sie den natürlichen Bewegungsablauf des einzelnen als auch die natürlichen Interaktionen zwischen mehreren Probanden.



**Abb. 12.8:** Softwarepakete zur Punktverfolgung (SIMI Reality Motion Systems GmbH 2011): (a) Medizinische Wirbelsäulenanalyse durch die Erfassung farbiger LED-Marker mit *Simi Aktisys*. (b) Automatische Verfolgung von auf einen Schwimmer aufgeklebten Zielmarken mit *Simi MatchiX*.



**Abb. 12.9:** *Motion Capturing Studio*: (a) Studioumgebung zur vollständigen Körper- und Bewegungserfassung (Vicon House of Moves, Inc. 2008). (b) NIR-Megapixelkamera zur Erfassung retroreflektierender Positionsmarker (Inition: Everything in 3D 2011).



**Abb. 12.10:** Bewegungserfassung in Filmindustrie und Sportwissenschaft (Motion Analysis Corporation 2011): (a) Softwareumgebung *Cortex* mit *Talon Plug-ins* zur echtzeitfähigen Verfolgung und Animation menschlicher Bewegungen. (b) Aufnahme, Rekonstruktion und Analyse von Golfabschlägen durch markerbasierte Verfahren der 3D-Bewegungsanalyse.

## 12.4 Diskussion

Unter den Aspekten einer weitgehend beobachterunabhängigen und möglichst automatisierten Datenerhebung zur Objektivitäts- und Reliabilitätserhöhung, Zeitersparnis, Entlastung personeller Ressourcen, Reduktion von Beobachtereinflüssen sowie Erhöhung des Auflösungsgrades kommen die oben diskutierten und zumeist kommerziell erhältlichen Verfahren der automatischen 3D-Bewegungsanalyse in Frage. Der sensorielle, algorithmische und rechentechnische Aufwand der Auswertung multiokularer Bildsequenzen ist aber insgesamt sehr hoch. Auch ergeben sich Einschränkungen der Anwendbarkeit für Experimente zum menschlichen Verhalten aufgrund der notwendigen Signalisierung.

Hier eignen sich 3D-Kameras als monosensorielle Systeme zur Echtzeit-Tiefenbildakquisition mit hoher zeitlicher Auflösung besonders für Anwendungen im breiten Gebiet der menschlichen Bewegungsanalyse (Kapitel 8). Ihr entscheidender Vorteil gegenüber multiokularen Systemen liegt im Wegfall des Arbeitsschritts der stereoskopischen Zuordnung. Auch lassen sie sich aufgrund ihrer geringen Abmaße leicht und unauffällig in einer entsprechenden Experimental- bzw. Arbeitsumgebung integrieren. Trotz der noch vergleichsweise geringen räumlichen Auflösung stellen 3D-Kameras perspektivisch ein leistungsfähiges Werkzeug der Datenakquisition für Aufgabenstellungen der berührungslosen Analyse des menschlichen Bewegungsverhaltens dar. Ihr Einsatz bedingt die Adaption, Weiter- und Neuentwicklung von Verfahren der Bildanalyse, was Gegenstand des folgenden Kapitels 13 ist.



## Analyse von 3D-Kamerabildsequenzen

Das folgende Kapitel präsentiert eine kleine Auswahl von seit Aufkommen der noch relativ jungen Tiefenbildtechnologie in der Literatur vorgestellten Ansätzen zur zwei- und dreidimensionalen Bewegungsanalyse aus 3D-Kamerabildsequenzen. Es werden unterschiedliche Methoden und daraus konzipierte Verfahren vorgestellt, welche sich primär im Bereich der Detektion, Lokalisierung und Interpretation menschlicher Bewegungen ansiedeln.

Der Autor kategorisiert (und bewertet) die gesichteten Arbeiten hauptsächlich im Hinblick auf die verwendeten 3D-Kamerainformationen: Die einfachste Möglichkeit der Auswertung von ToF-Bildsequenzen (*Time-of-Flight*) ist die Anwendung bekannter Methoden und Verfahren der Photogrammetrie und der Bildanalyse auf den Amplituden- oder den Entfernungskanal ([Abschnitt 13.1](#)). Dazu zählen auch solche Ansätze, die direkt mit einer über die inneren Orientierungsparameter der Kamera, die diskrete Sensorposition und die gemessene Schrägdistanz berechneten 3D-Punktwolke arbeiten. Die Verwendung von lediglich einem Informationskanal ist für eine Vielzahl von Anwendungen zweckmäßig, vor allem da einer der zwei Kanäle Tiefeninformationen in Videorate direkt bereitstellt. Sie nutzt aber nicht das gesamte Potential der ToF-Technologie aus. Deshalb ist die Einbeziehung aller zur Verfügung stehenden Beobachtungsarten in eine Berechnungsroutine naheliegend und wird von vielen Arbeitsgruppen in sequentieller Art vorgenommen ([Abschnitt 13.2](#)). Zunächst werden adaptierte oder neu entwickelte Methoden der Bewegungsextraktion unabhängig voneinander in beiden Kanälen angewendet und später zu einer Lösung kombiniert. Erfolgt darüber hinaus eine funktionale und im besten Fall auch stochastische Verknüpfung der Amplituden- und Distanzbeobachtungsarten in einem integrierten Ansatz, kann von einer simultanen Auswertung gesprochen werden ([Abschnitt 13.3](#)). Diese Vorgehensweise schöpft den Informationshaushalt vollständig aus und gilt somit als genau und zuverlässig.

Weitere Unterteilungen sind prinzipiell denkbar, bspw. entsprechend dem gewählten Berechnungsansatz (daten- oder modellbasiert, einfache Bildanalyseverfahren oder wahrscheinlichkeitsbasierte Ansätze), der Fusion mit zusätzlichen Sensoren (Integration in Stereo- oder Mehrkamerakonfigurationen) oder der Art der gewonnenen Bewegungsinformationen (einzelne Bewegungsvektoren, vollständige Flussfelder oder zeitlich aufgelöste Parametersätze komplexer (Mensch-)Modelle). Da in den letzten fünf Jahren Arbeitsgruppen verschiedener Wissenschaftsgebiete vielfältige Herangehensweise präsentierten, wird diese Unterteilung im Folgenden nicht vollständig vorgenommen.

In den Publikationen sind meistens nur die entwickelten Methoden sowie deren Implementierungen in lauffähige Verfahren beschrieben. Sie dokumentieren anschaulich das Potential der ToF-Technologie, ohne Aussagen über die Richtigkeit der Ansätze bereitzustellen. Insofern Informationen über Genauigkeit und Zuverlässigkeit vorliegen, werden diese entsprechend angegeben.

## 13.1 Auswertung des Amplituden- oder Entfernungskanals

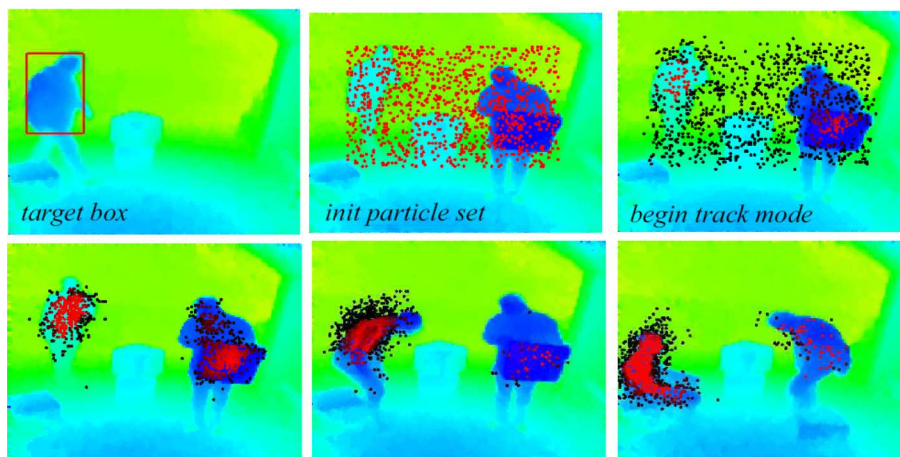
### 13.1.1 Amplitudenbildsequenzanalyse

Die folgenden zwei Beiträge verwenden nur die Amplitudenbilder zur Verfolgung von Bewegungen. Die Entfernungsinformation wird an der geschätzten Position interpoliert, nicht aber zur eigentlichen Berechnung herangezogen. In (Beyeler 2006) werden Infrarotreflektoren in Echtzeit am Beispiel einer medizinischen Anwendung zur spatiotemporalen Rekonstruktion von der Position und der Orientierung chirurgischer Instrumente verfolgt. Beyeler erachtet für einen Arbeitsbereich von 0,5 m bis 1 m die Eignung einer ToF-Kamera in Umgebungen der erweiterten Realität (*Augmented Reality*) für ausreichend, schließt jedoch invasive Eingriffe aufgrund des geringen Genauigkeitspotentials aus. Kohoutek (2008) detektiert einen mobilen Industrieroboter in Amplitudenbildsequenzen über Kantenfilteroperationen. Anschließend verfolgt er den Roboter durch Methoden des lokalen zweidimensionalen Optischen Flusses mit dem Ziel, den Arbeitsbereich dynamisch nachzuführen und zu überwachen (Abschnitt 7.5.1).

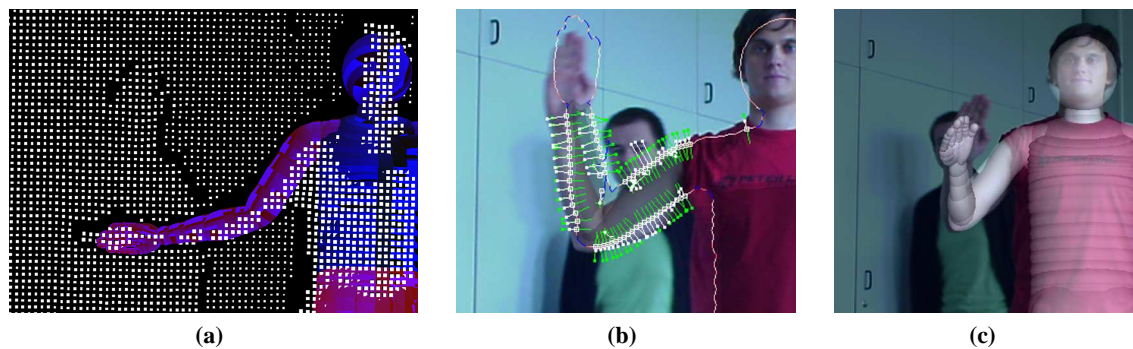
### 13.1.2 Entfernungsbildsequenzanalyse

Göktürk und Rafi (2005) beschreiben die Anwendung einer 3D-Kamera bei der Kontrolle der Besetzung von Sitzen in einem Fahrzeug, um über den Einsatz eines Airbags im Falle eines Unfalls zu entscheiden. Die Verfolgung der Köpfe von Fahrzeuginsassen beruht einzig auf der Grundlage von Entfernungsbildern. In einer Trainingsphase werden durch interaktive Selektion repräsentative Tiefensignaturen der Köpfe berechnet, welche in der sich anschließenden *Tracking*-Phase über Korrelationstechniken mit den Tiefenwerten des aktuellen Bilds verglichen werden. Die Fehlerrate bei der Klassifikation liegt bei unter 2%.

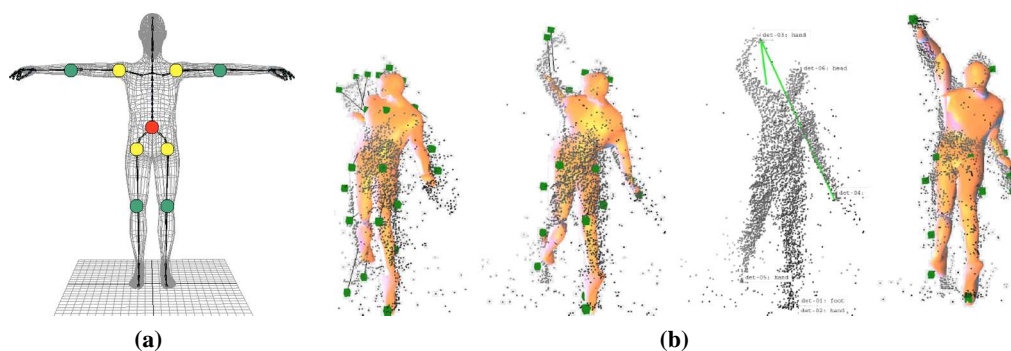
Kahlmann und Ingensand (2006) sowie Kahlmann u. a. (2007) setzen ToF-Kameras für Überwachungsaufgaben ein und detektieren sich bewegende Personen durch Hintergrundsubtraktion, Schwellenwertbildung und Clusteranalyse in Entfernungsbildern (Abbildung 13.1). Die segmentierten Daten bilden die Grundlage für einen wahrscheinlichkeitsbasierten *CONDENSATION-Tracker* (*Conditional Density Propagation for Visual Tracking*), welcher erstmals in (Isard & Blake 1998) vorgestellt und in (Koller-Meier 2000) für die Verfolgung mehrerer Personen in Tiefenbildsequenzen erweitert wird. Die Autoren geben eine Genauigkeit von ca. 1 cm für kalibrierte 3D-Kameradaten an. Auch Gudmundsson u. a. (2010) wenden diesen Partikelfilter auf zuvor aus Distanzbildern segmentierte 3D-Kameradaten an, um die Parameter eines Handmodells für jeden Zeitschritt bestimmen zu können. Die Genauigkeit und Zuverlässigkeit der Bildverarbeitungsprozedur wird lediglich visuell eingeschätzt. Ähnlich werden menschliche Bewegungen in (Ganapathi u. a. 2010) mit bis zu 10 Bildern pro Sekunde einzig aus den Entfernungsbildsequenzen einer 3D-Kamera rekonstruiert, wobei dieser Ansatz ebenfalls auf der Kombination eines aus 15 starren Körperteilen gebildeten Menschmodells mit einem wahrscheinlichkeitsbasierten Partikelfilter beruht (Abbildung 13.3). Ganapathi u. a. berichten jedoch von sehr großen Fehlern im dm-Bereich, welche aus einer ungenügenden bzw. gar



**Abb. 13.1:** Personendetektion in ToF-Bildsequenzen auf der Grundlage eines wahrscheinlichkeitsbasierten Partikelfilters (Kahlmann u. a. 2007).



**Abb. 13.2:** Die Verwendung von 3D-Tiefenbildern zur verbesserten Segmentierung bei der Bestimmung von Bewegungsinformationen durch Silhouettenansätze (Grest u. a. 2007): (a) Segmentiertes 3D-Kameratiefenbild. (b) Silhouettenkorrespondenzen. (c) Ein 2D-Kamerabild wird überlagert mit dem geschätzten Menschmodell.



**Abb. 13.3:** Visuelle Personenverfolgung mit einem Partikelfilter auf der Basis von 3D-Kameratiefenbildern (Ganapathi u. a. 2010): (a) Menschmodell. (b) Modellierte Pose eines Tennisspielers, überlagert mit den zuvor segmentierten und anschließend zur Bewegungsschätzung herangezogenen 3D-Kameradaten.

nicht durchgeführten Distanzkalibrierung resultieren können.

Knoop u. a. (2006) stellen einen auf 3D-Kameradaten basierenden Algorithmus zur dreidimensionalen Verfolgung menschlicher Körperbewegungen unter der Verwendung eines Stereokamerasystems und einer ToF-Kamera vor. Für das eingeführte *Tracking*-System wird ein 3D-Körpermodell, bestehend aus zehn Zylindern mit zusätzlichen Bedingungen zwischen den einzelnen geometrischen Primitiven eingeführt. Über ICP-basierte Methoden der Punktwolkenzuordnung (*Iterative Closest Point Algorithm*) wird die Körperhaltung für jedes Bild der Sequenz modelliert. Eine Einschätzung des Genauigkeits- und Zuverlässigkeitspotentials erfolgt über einen manuellen Vergleich der rekonstruierten Körperposen mit den aufgenommenen Videodaten. Auch die in (Breidt u. a. 2010) vorgestellte modellbasierte Rekonstruktion eines Gesichtes und die darauf aufbauende Interpretation der beobachteten Mimik beruht auf einem ICP-Algorithmus (Abbildung 13.4).

Grest u. a. (2007) und Gudmundsson u. a. (2008) kombinieren Entfernungsinformationen zur Verfolgung menschlicher Bewegungen aus Tiefenbildern einer 3D-Kamera und aus konventionellen RGB-Bildern einer bzw. mehrerer 2D-Videokameras (Abschnitt 6.2.4; Abbildung 13.2). Durch die Sensorfusion wird die Qualität der Personenverfolgung erhöht, vor allem aufgrund einer durch die Hinzunahme von 3D-Kamerainformationen verbesserten Segmentierung der verfolgten Person und des sich teilweise ändernden Hintergrunds. Die Orientierungsparameter des Körpers berechnen die Arbeitsgruppen anschließend über eine ausgleichende Transformation der segmentierten Daten auf Mensch- und Bewegungsmodelle. In beiden Publikationen werden Genauigkeit und Zuverlässigkeit nur formal diskutiert, aber nicht anhand konkreter Versuchsreihen belegt.

Breuer u. a. (2007) extrahieren in einer Pilotstudie Handbewegungen (Position und Orientierung) aus 3D-Kamerasequenzen über Schwellenwertoperationen und Segmentierung mit anschließender Berechnung der Hauptkomponenten und Einpassung der Punktwolken in vorliegende Handmodelle. Sie dokumentieren die Qualität des Algorithmus als *gut*, berichten aber auch von Fehlern bei der Modelleinpassung, induziert durch zu schnelle Bewegungen.

In (Diraco u. a. 2010) wird ein Verfahren zur automatischen Detektion von Stürzen älterer Menschen im Haushalt vorgestellt, welches dynamische Regionen in 3D-ToF-Punktwolken in Echtzeit über einen Bayesschen Filter segmentiert (Abbildung 13.5). Die Körperpose wird zuerst über den Schwerpunkt der Punktwolke (Translation) und anschließend, nach der Extraktion eines abstrahierten 3D-Knochengengerüsts, über die Orientierung des Rückens (Rotation) rekonstruiert. Der zeitlich aufgelöste Abstand zwischen dem Köperschwerpunkt und dem als eine Ebene modellierten Boden gilt als Maß für die Alarmauslösung im Falle eines Sturzes.

## 13.2 Sequentielle Auswertung des Amplituden- und Entfernungskanals

Reiser und Kubacki (2007) beschreiben erste Versuche zur Integration einer 3D-Kamera bei der Sicherheitsüberwachung des Arbeitsraums eines Industrieroboters. Die Robotergelenke werden durch Kugeln simuliert, die in einer 3D-Kamerabildsequenz detektiert und verfolgt werden sollen. Nach einer Hintergrundsegmentierung und einer Datenglättung werden die Primitive über Kantenfilterung



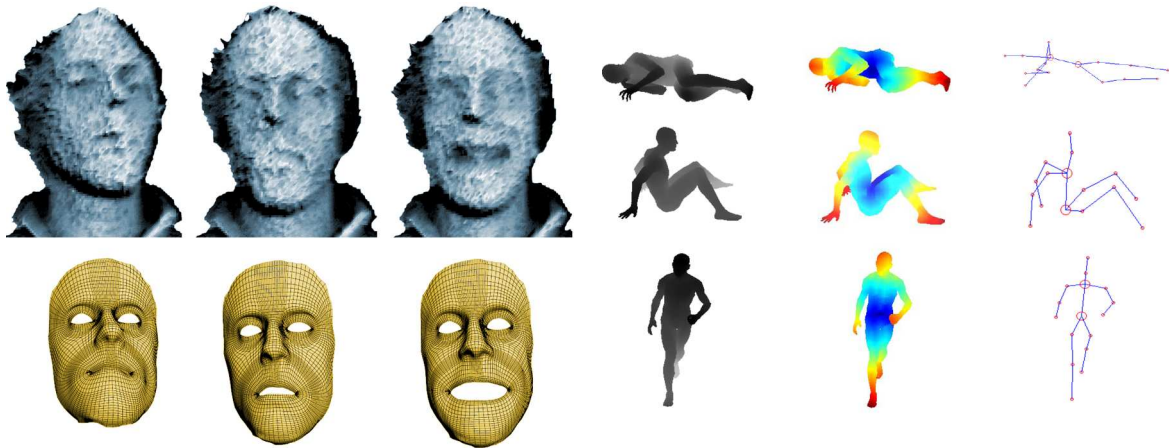
und Ellipsenfitting unter sequentieller Verwendung beider Kanäle extrahiert, die Kugelmittelpunktkoordinaten in den Objektraum transformiert und die Verschiebung zum vorherigen Zeitpunkt gespeichert. Der relative Positionierungsfehler beträgt wenige Millimeter.

Meers und Ward (2008) bestimmen Kopfpositionen in Echtzeit. Das geschieht zunächst annähernd durch eine einfache Profilanalyse im Tiefenbild. Anschließend wird die Nase als einzigartiges und zu meist nicht verdecktes Merkmal im Amplitudenkanal über bestimmte Reflexionseigenschaften sowie im Distanzkanal über die Analyse von Oberflächennormalen detektiert und als Ausgangspunkt für eine sphärische Profilanalyse im Distanzkanal genutzt. Die genaue Kopfrekonstruktion pro Zeitschritt erfolgt zunächst prototypisch über Profilzuordnungen. Im Laufe der Sequenzanalyse vermaschen Meers und Ward die jeweiligen Rekonstruktionen zu einem 360°-Kopfmodell.

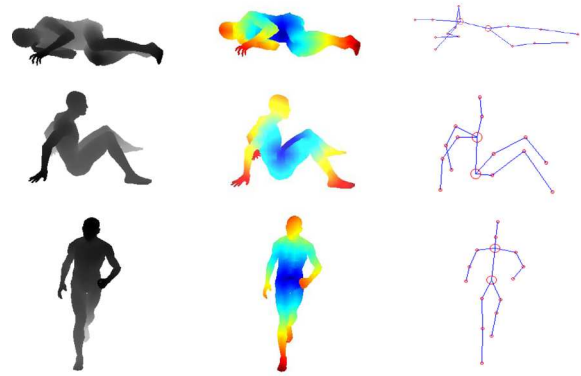
In (Beuter u. a. 2010; Swadzba u. a. 2010) werden am Beispiel einer Anwendung der mobilen Haushaltsrobotik alle freibeweglichen Objekte wie Stühle oder Türen, sich eigenständig bewegende Personen sowie die statische Hintergrundszene kartiert (Abbildung 13.6). Die Daten beider Kanäle werden von Ausreißern bereinigt, geglättet und grob über eine Schwellenwertbildung segmentiert. Der Amplitudenkanal bildet die Grundlage für die Berechnung eines Bewegungsvektorfelds durch den lokalen zweidimensionalen Optischen Fluss (Abschnitt 7.5.1). Die Distanzbeobachtungen werden zusammen mit den inneren Orientierungsparametern des Sensors zur Berechnung zeitaufgelöster 3D-Punktwolken herangezogen. Abschließend erfolgt ein 6D-Clustering auf der Basis aller vorhandenen Attribute – also den Amplituden- und Distanzwerten, den ermittelten Geschwindigkeiten und den 3D-Koordinaten – sowie die Einpassung einfacher zylindrischer Primitive. Die Restklaffungen der 3D-Datenpunkte zum geschätzten Modell (hier über euklidische Distanzen berechnet), betragen im Mittel 2 cm.

Auch in (Holte u. a. 2010) werden zur Gestendetektion zweidimensionale optische Flussfelder aus Amplitudenbildsequenzen geschätzt und die so erhaltenen Geschwindigkeitsvektoren den transformierten 3D-Objektkoordinaten zugeordnet (Abbildung 13.7). Unter der Verwendung dieser spatio-temporalen Information können 22 einfache Gesten mit einer Erfolgsrate von bis zu 94% erkannt werden.

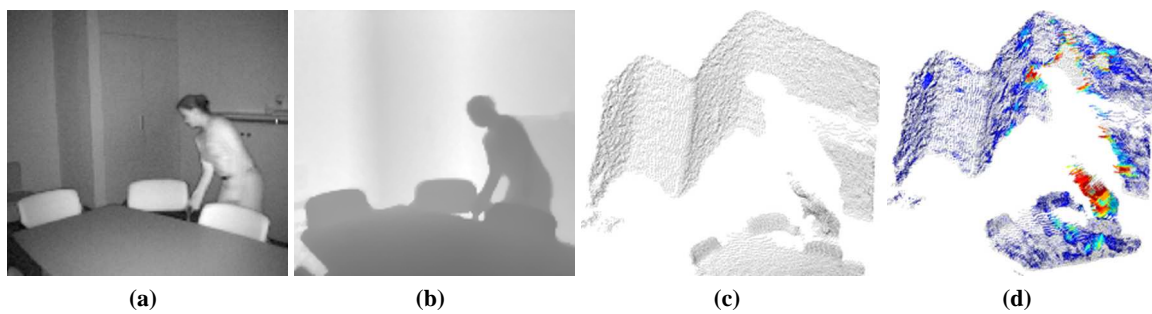
In (Lahamy & Lichti 2011) werden die Interaktionsmöglichkeiten zwischen Mensch und Maschine durch den Einsatz einer 3D-Kamera erweitert. Die Implementierung einer einfachen Variante des *Mean Shifts* (Cheng 1995) erlaubt die robuste echtzeitfähige Handverfolgung. Lahamy und Lichti segmentieren initial die Hand in beiden Kanälen und verfolgen sodann 3D-Handsegmente innerhalb eines  $(20\text{ cm})^3$  großen Kuboids in definierten Merkmalsräumen ohne die Annahme einer Verteilungsfunktion. Die mittlere quadratische Abweichung der berechneten Schwerpunktkoordinaten wird mit 1 cm angegeben. Die Arbeitsgruppe stellt keinen entfernungs- oder integrationszeitabhängigen Genauigkeitsverlust fest, berichtet aber von Problemen bei Handbewegungen schneller als  $10\text{ cm/s}$ .



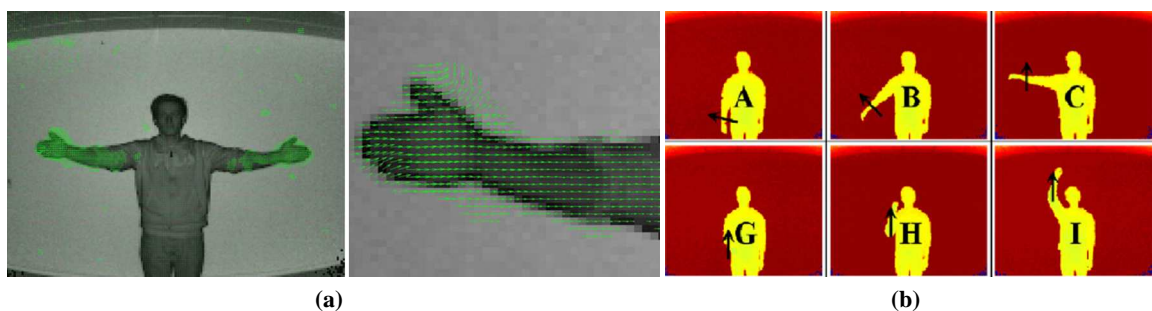
**Abb. 13.4:** Ergebnisse einer modellbasierten Gesichtrekonstruktion (nach Breidt u. a. 2010). Oben: 3D-Kamerapunktwolke. Unten: Rekonstruiertes Modell.



**Abb. 13.5:** Liegen, Sitzen, Stehen sind neben Bücken die drei Hauptkörperstellungen, welche automatisch erfasst und interpretiert werden können (nach Diraco u. a. 2010). Links, Mitte: Tiefenwertdarstellung. Rechts: Abstrahiertes Knochengeriüst.



**Abb. 13.6:** Bewegungsanalysen in 3D-Kamerabildsequenzen zur Umgebungskartierung (Swadzba u. a. 2010): (a) Amplitudenbild. (b) Entfernungsbild. (c) Ausreißerbereinigte und geglättete 3D-Punktwolke. (d) 3D-Punktwolke mit farbkodierten Geschwindigkeitsvektoren, geschätzt durch Methoden des Optischen Flusses.



**Abb. 13.7:** Gestendetektion und -interpretation auf Basis von 3D-Kameradaten (Holte u. a. 2010): (a) Amplitudenbild mit überlagerten 2D optischen Flussvektoren. (b) 6 von 22 möglichen Gesten, dargestellt in farbkodierten Entfernungsbildern, die automatisch interpretiert werden können.

### 13.3 Simultane Auswertung des Amplituden- und Entfernungskanals

Jensen u. a. (2009) untersuchen menschliche Bewegungen auf einem Laufband (Abbildung 13.8). Der vorgestellte Ansatz segmentiert automatisch den menschlichen Körper und analysiert den Gangzyklus unter simultaner Verwendung des Intensitäts- und Entfernungskanals. Die Einpassung der 3D-Kameradaten in ein abstrahiertes 2D-Menschmodell wird als Minimierungsproblem formuliert, wobei eine aus vier speziell für diese Anwendung aufgestellten Termen zusammengesetzte additive Zielfunktion die Integration beider Kanäle über einzuhaltende Nachbarschaftskonsistenzen, a-priori Wissen zur Körperoberfläche und Histogrammanalysen vornimmt.

Darüber hinaus adaptieren M. Schmidt u. a. (2008) sowie Westfeld und Hempel (2008) aus der Bildanalyse bekannte Zuordnungsverfahren, nämlich den Optischen Fluss sowie die Bildzuordnung nach der Methode der Kleinsten Quadrate (Abschnitte 7.5 und 7.6), und entwickeln erstmals unter simultaner Verwendung von Intensitäts- und Entfernungsdaten integrierte flächenbasierte spatiale und temporale Verfahren zur Objektzuordnung und -verfolgung in 3D-Kamerabildsequenzen. Im Folgenden werden die beiden Vorgehensweisen im Hinblick auf das im Kapitel 14 behandelte integrierte 2,5D *Least Squares Tracking* (LST) näher vorgestellt:

#### 13.3.1 Range Flow

Spies u. a. (2002) erfassen Blattbewegungen dreidimensional durch Verfahren der strukturierten Beleuchtung (Abschnitt 6.2.1). Das resultierende Geschwindigkeitsvektorfeld  $f$ , das die Bewegung der Objektfläche über eine Funktion  $Z(X, Y, t)$  beschreibt, heißt *Range Flow* (Yamamoto u. a. 1993) oder *Scene Flow*, wenn dieser aus Stereo- oder Mehrbildaufnahmen bestimmt worden ist (Zhang & Kambhamettu 2000). In Analogie zur *Brightness Constancy Constraint Equation* (BCCE) des differentiellen zweidimensionalen Optischen Flusses (Abschnitt 7.5) kann für kontinuierlich verlaufende 2,5D-Flächendaten eine *Range Flow Constraint Equation* (RFCE) formuliert werden:

$$\frac{dZ}{dt} = \frac{\partial(Z, Y)}{\partial(x', y')} U + \frac{\partial(X, Z)}{\partial(x', y')} V + \frac{\partial(Y, X)}{\partial(x', y')} W + \frac{\partial(X, Y, Z)}{\partial(x', y', t)} = 0 \quad (13.1)$$

mit

$$f = \begin{bmatrix} U & V & W \end{bmatrix}^T = \frac{d}{dt} \begin{bmatrix} X & Y & Z \end{bmatrix}^T \quad \text{Geschwindigkeit des Tiefenflussfelds}$$

Zusätzlich liefern optische Tiefenbildsensoren Intensitätswerte, die Spies u. a. zur Aufstellung der um die Lateralkomponenten  $(U, V)$  des Geschwindigkeitsvektorfelds  $f$  erweiterten BCCE (7.9) nutzen und so beide geometrischen Modelle verknüpfen:

$$\frac{dI}{dt} = \frac{\partial(I, Y)}{\partial(x', y')} U + \frac{\partial(X, I)}{\partial(x', y')} V + \frac{\partial(X, Y, I)}{\partial(x', y', t)} = 0 \quad (13.2)$$

In seiner Dissertationsarbeit wendet M. O. Schmidt (2008) diese Erweiterung des Optischen Flusses auf 3D-Kamerabildsequenzen an. Er substituiert den Objektkoordinatenvektor  $\mathbf{X}(X, Y, Z)$  durch

$\frac{\mathbf{x}'}{\|\mathbf{x}'\|}D$ , skaliert also den normierten Richtungsvektor  $\|\mathbf{x}'\| = \left\| \begin{bmatrix} x' & y' & -c \end{bmatrix}^T \right\|$  (ohne Berücksichtigung der Hauptpunktlage und möglicher Abbildungsfehler) mit der gemessenen Distanz  $D(\mathbf{x}', t)$  entsprechend dem geometrischen Modell einer ToF-Kamera (Abschnitt 8.3), und erhält:

$$\frac{dD}{dt} = U \left( Dx' - D_{x'} \|\mathbf{x}'\|^2 \right) + V \left( Dy' - D_{y'} \|\mathbf{x}'\|^2 \right) + Wd - DD_t \|\mathbf{x}'\| = 0 \quad (13.3)$$

mit

$$d = c \cdot D + \frac{\|\mathbf{x}'\|^2 (D_{x'} x' + D_{y'} y')}{c}$$

Darüber hinaus passt M. O. Schmidt die BCCE der aktiven Beleuchtung einer 3D-Kamera an und kombiniert schließlich beide Flussgleichungen in einem integrierten Ansatz nach der Methode der kleinsten Quadrate (MKQ), um die drei Translations- bzw. Geschwindigkeitskomponenten zu schätzen (Abbildung 13.9).

### 13.3.2 Integriertes Least Squares Tracking

Das in der Photogrammetrie zum Zwecke der Bildpunktzuordnung verbreitete *Least Squares Matching* (LSM; Abschnitt 7.6) wird in (Maas 2000) zur Auswertung von 2,5D-Flugzeuglaserscannerdaten adaptiert (vgl. Abschnitt 7.6.3). Maas transformiert dabei die in einer unregelmäßigen Dreiecksvermaschung gespeicherten Höhendaten benachbarter ALS-Streifen (*Airborne Laserscanner*) durch die Schätzung der drei Translationskomponenten aufeinander. Wie jedes differentiell arbeitende Bildzuordnungsverfahren benötigt auch dieser Ansatz signifikanten Signalkontrast, in diesem Fall also Höhengradienten. Diese fallen aber vor allem bei nahezu ebenen Geländeoberflächen wie bspw. Flachdächern sehr gering aus. In Konsequenz erweitert Maas (2001) die Methode um zusätzlich vorliegende ALS-Intensitätsbeobachtungen, welche den Zuordnungsprozess stützen sollen. Das vorgestellte Verfahren arbeitet sequentiell und bestimmt die Verschiebungskomponente orthogonal zur Flugrichtung aus den Höhendaten und die planimetrischen Variationen konventionell aus den Intensitätsdaten.

Darauf aufbauend stellt Westfeld (2007b) einen integrierten Ansatz eines *Least Squares Tracking* (LST) vor und verfolgt Oberflächensegmente in Bildsequenzen von 3D-Kameras auf der Basis von Intensitäts- und Tiefenbildern. Im Hinblick auf die Natur der 3D-Kameradaten wird eine geschlossene Lösung für die Zuordnung auf der Zeitachse vorgestellt und somit der Informationshaushalt durch die Verknüpfung der geometrisch-radiometrischen Modelle vollständig ausgeschöpft. Westfeld und Hempel (2008) erweitern den Ansatz um eine *Varianzkomponentenschätzung* (VKS) zur Bestimmung optimaler Gewichte der beiden Beobachtungsklassen. Sie führen außerdem eine robuste Berechnung der Varianz-Kovarianzmatrix ein. Die Standardabweichungen der lateralen Translationskomponenten im Bildraum betragen maximal  $1/25$  px, die Standardabweichung der Verschiebungskomponente orthogonal zur Bildebene ist entfernungsabhängig und beläuft sich auf 0,25% der gemessenen Distanz. Mit dem Ansatz werden Methoden zur Bestimmung der interpersonalen

Distanz, Körperorientierung und Kopfneigung interagierender Personen diskutiert und anhand von Versuchsreihen evaluiert ([Abbildung 13.10](#)).

## 13.4 Diskussion

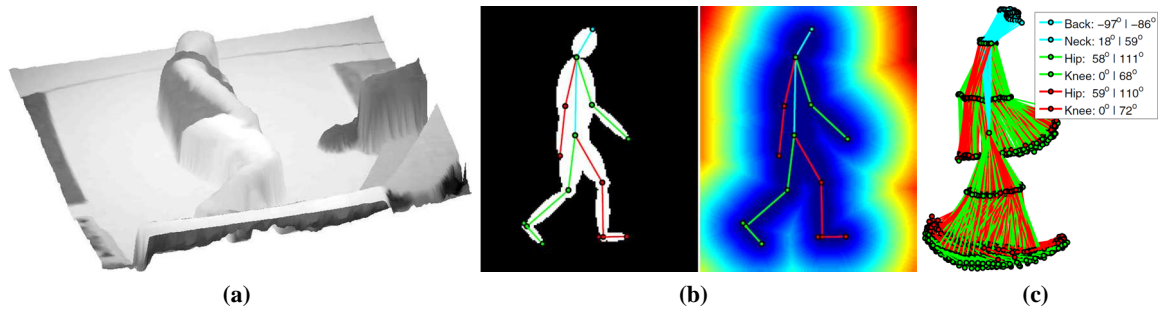
Zahlreiche wissenschaftliche Arbeitsgruppen beschäftigen sich mit der Detektion und Analyse menschlicher Bewegungsvorgänge und zeigen das Potential der Auswertung von 3D-Kamerabildsequenzen. Die oben aufgeführten Beispiele beruhen zum einen auf einfachen bildanalytischen und meist pixelbasierten Verfahren wie Schwellenwertbildung, Schwerpunktberechnung oder Clustering, zum anderen auf komplexen und in der Regel wahrscheinkeitsbasierten Ansätzen, die einen statisch-spatialen Zustand geometrisch über z. B. Transformationsvorschriften oder ganze 3D-Körpermodelle modellieren und/oder Zustände dynamischer Prozesse über Bewegungsmodelle wie den Kalman-Filter oder den [CONDENSATION-Tracker](#) schätzen.

Bedingt durch die Fülle unterschiedlicher Anwendungsschwerpunkte ist eine Einschätzung der Algorithmen lediglich anhand ihrer Komplexität nach Meinung des Autors nicht angebracht. Vielmehr wird die analytisch-konzeptionelle Vorgehensweise durch das zu erreichende Ziel bzw. durch die zur Analyse notwendigen Messdaten bestimmt. Für eine Aufgabenstellung ist bspw. die Abstraktion des gesamten Menschen über ein Körpergerüst notwendig, eine andere hingegen benötigt den Bewegungsverlauf eines ganz bestimmten Körperpunkts über die Zeit. Wünschenswert und sinnvoll erscheint aus geodätisch-photogrammetrischer Sicht die Verarbeitung von 3D-Kameradaten in integrierten Verfahren unter simultaner Verwendung aller zur Verfügung stehenden Beobachtungsarten sowie die adäquate Gewichtung der unterschiedlich genauen Beobachtungsgruppen als auch die Möglichkeit zur Ableitung statistischer Genauigkeitsangaben.

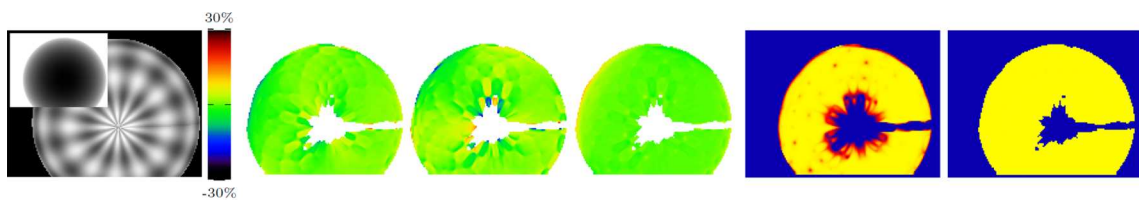
In diesem Kontext sind zwei bekannte Verfahren zur Bewegungsschätzung in ToF-Bildsequenzen weiterentwickelt worden: Der in der *Computer Vision* beheimatete Optische Fluss wird in ([M. Schmidt u. a. 2008](#)) um die Tiefeninformationen zum *Range Flow* erweitert. Das aus der Photogrammetrie bekannte 2D-LSM adaptiert [Westfeld \(2007b\)](#) auf die Daten einer 3D-Kamera zu einem 2,5D-LST. Beide Ansätze verknüpfen Amplituden- und Entfernungsinformationen durch einen gemeinsamen Transformationsparametersatz. Sie konstatieren, dass sich die Bewegungsschätzung unter der Ausnutzung des gesamten Informationshaushalts robuster gestaltet, da bspw. ein in einem Kanal aufgrund auftretender Korrespondenzprobleme schlecht konditioniertes Gleichungssystem durch die teilweise komplementäre Information des zweiten Kanals gestützt wird.

Die oben angeführten Publikationen zum integrierten LST entstanden begleitend während der Entwicklung des im Rahmen der vorliegenden Arbeit präsentierten spatialen und temporalen Zuordnungsverfahrens 2,5D-LST. Eine detaillierte Darstellung des Prinzips mit der Vorstellung eines erweiterten Parametersatzes sowie die Validierung der Methode folgt im [Kapitel 14](#).

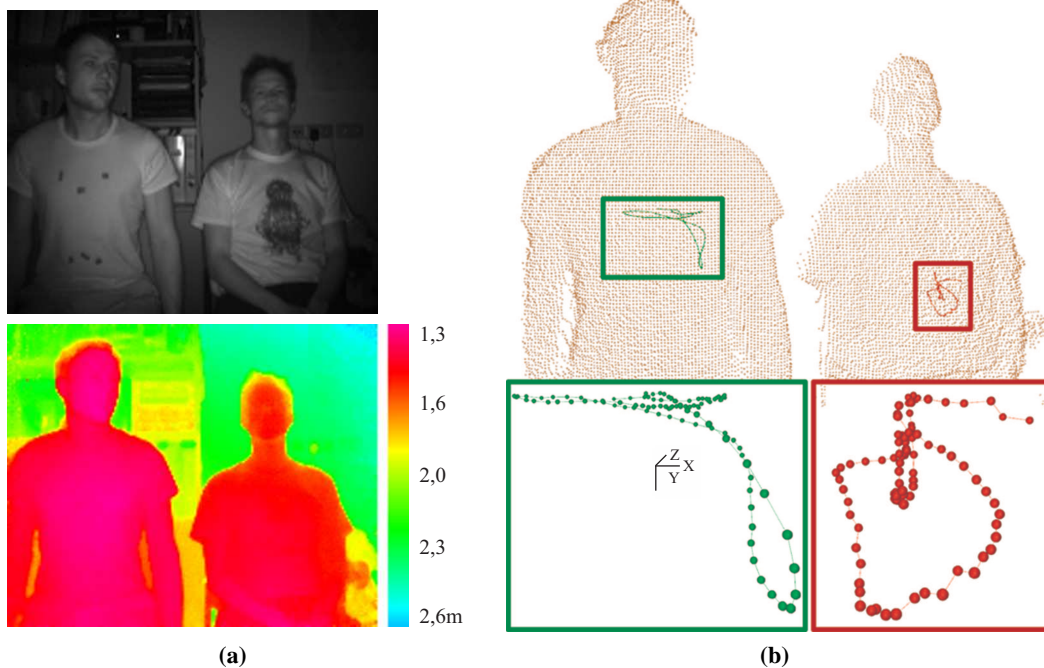




**Abb. 13.8:** Ganganalysen am Laufband auf Basis von 3D-Kameradaten (Jensen u. a. 2009): (a) 3D-Amplitudenbild. (b) Rastermodell und das korrespondierende farbkodierte Entfernungsbild. (c) Visualisierung der Gangzyklen.



**Abb. 13.9:** Visualisierung des *Range Flows* einer Kugel mit einem Radius von 50cm in einer Entfernung von 2,2 m (nach M. O. Schmidt 2008). Zusätzlich untersucht wurden in diesem Zusammenhang auftretende Bewegungsartefakte zu jedem Abtastzeitpunkt (vgl. Abschnitt 9.8).



**Abb. 13.10:** Visualisierung der Eingabedaten und der Ergebnisse eines 2,5D-LST: (a) Amplitudenbild und farbkodiertes Entfernungsbild zweier Probanden. (b) Die dazu korrespondierende 3D-Punktwolke, überlagert mit den berechneten Trajektorien der Körperschwerpunkte.

## Integriertes Least Squares Tracking

Die Arbeit beschäftigt sich mit der Auswertung zeitaufgelöster 3D-Kameradaten mit dem Ziel, Bewegungsvektoren bzw. Bewegungsvektorfelder zu extrahieren (Kapitel 7). Die Entwicklung der hierfür notwendigen bildanalytischen Methode ist Gegenstand des folgenden Kapitels.

Im Gegensatz zu konventionellen Kameras nehmen 3D-Kameras zusätzlich zum Amplitudenbild simultan ein Entfernungsbild auf (Kapitel 8). In logischer Schlussfolge können die ermittelten Schrägdistanzen bzw. die reduzierten Horizontaldistanzen mit ihren komplementären Grauwertinformationen in einen gemeinsamen funktionalen und stochastischen Zusammenhang gebracht und simultan zur Berechnung von optischen Flussfeldern in 3D-Kamerabildsequenzen herangezogen werden. Dabei soll die zu entwickelnde Methode die folgenden und bereits im Kapitel 2 benannten Merkmale aufweisen:

- Simultaneität: Eine geschlossene Formulierung der Grauwert- und Entfernungsbeobachtungen in einem integrierten geometrischen und stochastischen Modell ist wünschenswert, um das vollständige Informationspotential auszunutzen.
- Adaptabilität: Die Erweiterung des Parametersatzes des verwendeten geometrischen Modells sowie die Integration von anwendungsspezifischem a-priori Wissen in Form von Bedingungs-gleichungen soll gewährleistet sein.
- Genauigkeit: Eine verlässliche subpixel- bzw. submillimetergenaue Messung von Verschiebungen im Bildraum bzw. in Aufnahme-richtung wird angestrebt.
- Noninvasivität: Die zu entwickelnde Methode soll auf der Grundlage berührungslos aufgenommener 3D-Kameradaten basieren, um Eingriffe in die Experimentalumgebungen seitens des Vermessers zu vermeiden.
- Validierung: Informationen über die Genauigkeit und Zuverlässigkeit der Bewegungsschätzung als Entscheidungskriterien für die spätere Integration in eine Verarbeitungskette sowie Aussagen über das Genauigkeitspotential der Eingangsdaten sollen inhärent zur Verfügung gestellt werden.

### 14.1 Übersicht

Im Abschnitt 13.3 wurde bereits angedeutet, dass es sich bei der im Rahmen der vorliegenden Arbeit entwickelten Methode zur 3D-Kamerabildzuordnung um einen Ansatz nach der Methode der kleinsten Quadrate (MKQ; Abschnitt 4.4.1) handelt, welcher in seiner ursprünglichen Form bereits seit Mitte der 1980er Jahre in der Photogrammetrie bekannt ist (*Least Squares Matching*,

LSM; Abschnitt 7.6). Im Gegensatz zu dem in der *Computer Vision* verbreiteten Optischen Fluss (Abschnitt 7.5), welcher lediglich zwei laterale Verschiebungsparameter bestimmt, bzw. zu dem um die Tiefenkomponente erweiterten *Range Flow* (Abschnitt 13.3), liegt der Bildzuordnung nach der Methode der kleinsten Quadrate eine höher parametrisierte Affintransformation zugrunde (Abschnitt 4.1). Es können so nicht nur reine Translationen, sondern auch Rotationen und Maßstabsänderungen des Objekts modelliert werden; in der um die Distanzmessung erweiterten 2,5D-Form auch Verschiebungen in und Neigungen orthogonal zur Aufnahmerichtung. Darüber hinaus ist eine für die automatische Aufstellung des stochastischen Modells genutzte *Varianzkomponentenschätzung* (VKS) problemlos in die MKQ-Berechnungsroutine implementierbar (Abschnitte 4.5 und 4.6). Auch besteht die Möglichkeit, anwendungsspezifisch definierte Zwangsbedingungen flexibel in den Zuordnungsprozess zu integrieren (Abschnitt 4.4.3). Bei menschlichen Bewegungsanalysen werden solche Restriktionen bspw. aus dem zugrunde liegenden parametrisierten Menschmodell (räumliche Modellannahme) und dem zu erwartenden Bewegungsablauf (zeitliche Modellannahme) abgeleitet.

Die einfache Anwendung von 2D-LSM auf die Grauwertbilder einer 3D-Kamera ist problemlos möglich und resultiert in 2D-Bewegungsvektorfeldern (Abschnitt 7.6.1). Die Tiefenbildinformation bleibt bei dieser Vorgehensweise aber unberücksichtigt. Als Erweiterung können Entfernungsbilder als zweiter Kanal in ein 2D-LSM eingehen und gemeinsam mit den Intensitätsdaten auf der Basis einer einfachen 2D-Affintransformation ausgewertet werden. Der in (Maas 2001) gezeigte Ansatz kombiniert so in Abhängigkeit der Kontrastverhältnisse radiometrische und geometrische Oberflächenmessdaten in einem LSM-Ansatz. Ähnlich wie die simultane Verwendung aller drei Farbkäle einer konventionellen RGB-Kamera zeigt diese Strategie ein robustes Verhalten bei Texturschwäche in einem der beiden Kanäle, nutzt aber dennoch nicht die Tiefeninformation als geometrische Größe aus. Ein Übergang auf die Voxeldatenstruktur eines 3D-LSM ist ausgeschlossen (Abschnitt 7.6.2), da *Time-of-Flight*-Sensoren (ToF) lediglich Oberflächen- und keine Volumendaten erzeugen. Das aus der Literatur bekannte *Least Squares 3D Surface Matching* (LS3D; Grün & Akca 2004) kann die geometrischen Oberflächeninformationen zwar optimal verarbeiten (Abschnitt 7.6.3), nutzt aber wiederum nicht die Intensitätsinformation der Amplitudenmessung.

Basierend auf Maas' (2001) Arbeiten wird im Folgenden ein auf die Datenstruktur einer 3D-Kamera erweiterter LSM-Lösungsansatz zur Verfolgung kleiner Objektflächensegmente vorgestellt, welcher entsprechend den in der Einleitung des Kapitels aufgestellten Zielen als geschlossener Ansatz unter simultaner Verwendung aller vorliegenden Beobachtungen formuliert wird.

## 14.2 Funktionales Modell

Die um das interessierende Bildelement  $(x', y')$  zentrierte Nachbarschaft  $g_1$  kann affin in  $g_2$  transformiert werden. Eine ebene Affintransformation (4.2) bzw. (7.18) modelliert dabei Translationen, Rotationen und Maßstabsänderungen in jeweils beiden Koordinatenrichtungen und transformiert die Grauwerte des Musterfensters  $g_1$  in das Suchfenster  $g_2$  (vgl. auch Abschnitt 7.6 und Abbildung 7.9). Der vorliegende Ansatz wird darüber hinaus um zwei Koeffizienten  $(a_3, b_3)$  erweitert, welche mögli-



che Neigungen um die  $x'$ - bzw.  $y'$ -Achse berücksichtigen (Abbildung 14.1):

$$\begin{aligned}x'_2 &= x'_1 + \Delta x' = a_0 + a_1 x'_1 + a_2 y'_1 + a_3 x'_1 y'_1 \\y'_2 &= y'_1 + \Delta y' = b_0 + b_1 x'_1 + b_2 y'_1 + b_3 x'_1 y'_1\end{aligned}\quad (14.1)$$

Diese zwei zusätzlichen Parameter werden in der Aerotriangulation zur Modellierung systematischer Bildblockfehler angesetzt (Ebner 1976; Grün 1986). Die Notwendigkeit für den vorliegenden 2,5D-Ansatz wird im Folgenden verdeutlicht.

Der Abschnitt 14.2.1 stellt die ursprüngliche und linearisierte Beobachtungsgleichung des Amplitudenkanals auf. In einem zweiten Abschnitt 14.2.2 geschieht dies ebenfalls für den simultan aufgenommenen Entfernungskanal, wobei zusätzlich mögliche Abhängigkeiten der Distanzanpassung vom erweiterten Affintransformationsparametersatz aufgezeigt und korrigiert werden. Die Aufstellung der Koeffizientenmatrix und des Beobachtungsvektors erfolgt in den Abschnitten 14.2.4 und 14.2.5.

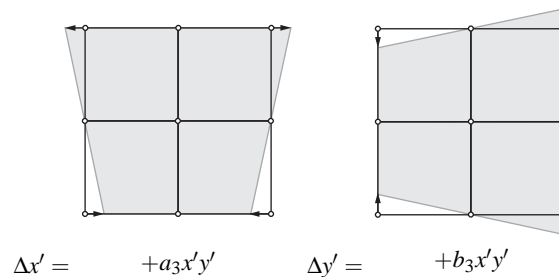


Abb. 14.1: Zusätzliche Neigungsparameter und ihre Auswirkungen im Bild (nach Kraus 1996).

### 14.2.1 Amplitudenkanal

Die Amplitudenwerte einer 3D-Kamera können in dem entwickelten 2,5D-LST in Analogie zu den in einer konventionellen Bildzuordnung nach der Methode der kleinsten Quadrate genutzten Intensitätswerten einer 2D-Kamera eingeführt werden (Abschnitt 7.6). Ein Musterfenster  $g_1^A$  und ein Suchfenster  $g_2^A$  mit einer identischen Größe von  $(R \times C)$  px, extrahiert aus zwei konsekutiven und in einem Vorverarbeitungsschritt um die Abbildungsfehler<sup>7</sup> korrigierten Amplitudenbildern einer 3D-Kamerabildsequenz, stellen die Beobachtungen für jede Bildposition  $\mathbf{x}'$  im Amplitudenkanal zur Verfügung. Die geometrische und radiometrische Relation zwischen diesen kleinen Bildausschnitten

<sup>7</sup> Die Berücksichtigung sensorpositionsabhängiger Abbildungsfehler ist bei der Kleinsten-Quadrate-Anpassung in der Regel nur notwendig, wenn sich die wahrscheinliche Bildposition signifikant änderte. Jedoch müssen zur Aufstellung der im folgenden Abschnitt 14.2.2 eingeführten Beobachtungsgleichungen des Entfernungskanals die gemessenen Schrägdistanzen  $D$  auf Horizontalabstände  $D_{\perp}$  unter der Verwendung der im Kapitel 11 berechneten Korrekturparameter reduziert werden (Abschnitt 8.3). Da der den Sensorpositionseinfluss beschreibende Distanzkorrekturterm  $\Delta D_{x'y'}$  bereits unter der Berücksichtigung von der Hauptpunktlage und der Objektivverzeichnung ermittelt wurde (Abschnitt 9.6), ist der Amplitudenkanal ebenfalls im Vorfeld durch bspw. indirektes Resampling um auftretende Abbildungsfehler zu korrigieren.

wird über die ursprüngliche Beobachtungsgleichung (7.17) formuliert:

$$g_1^A(x', y') - e^A(x', y') = r_0 + r_1 \cdot g_2^A(x', y') \quad (14.2)$$

In ihrer linearisierten Form lautet sie:

$$\begin{aligned} g_1^A(x', y') - e^A(x', y') = & (g_2^A)^0(x', y') + g_{x'}^A \cdot da_0 + g_{x'}^A \cdot x' \cdot da_1 + g_{x'}^A \cdot y' \cdot da_2 + g_{x'}^A \cdot x' \cdot y' \cdot da_3 \\ & + g_{y'}^A \cdot db_0 + g_{y'}^A \cdot x' \cdot db_1 + g_{y'}^A \cdot y' \cdot db_2 + g_{y'}^A \cdot x' \cdot y' \cdot db_3 \\ & + r_0 + r_1 \cdot (g_2^A)^0(x', y') \end{aligned} \quad (14.3)$$

mit

$$g_{x'}^A = \partial(g_2^A)^0(x', y') / \partial x', \quad g_{y'}^A = \partial(g_2^A)^0(x', y') / \partial y'$$

Die Bereitstellung der zur Linearisierung notwendigen Näherungswerte für die Translations- und Rotationsparameter ( $a_0, a_2, b_0, b_1$ ) wurde bereits im theoretischen [Abschnitt 7.6.1](#) diskutiert. Die Initialisierung der verbleibenden Parameter wird im folgenden [Abschnitt 14.2.2](#) durchgeführt.

Wie in ([Baltsavias 1991](#)) vorgeschlagen, werden die radiometrischen Korrekturterme  $r_0$  und  $r_1$  vor der eigentlichen Ausgleichung bestimmt und die Amplitudenwerte entsprechend korrigiert. Dies gewährleistet eine robustere Schätzung der verbleibenden geometrischen Parameter. Die radiometrischen Unterschiede im Amplitudenkanal werden bei einer aktiven Beleuchtung mit **nahem Infrarot** bspw. aufgrund der Abhängigkeit der Intensitätswerte von der Entfernung zum Aufnahmeobjekt oder durch weiterer das **Signal-Rausch-Verhältnis (SNR)** beeinflussende Quellen verursacht ([M. Hempel 2007](#)).

### 14.2.2 Entfernungskanal

Als zusätzlicher Beobachtungskanal stehen die in den Entfernungsbildausschnitten  $g_1^{D_\perp}$  bzw.  $g_2^{D_\perp}$  gespeicherten und bereits in einem Vorverarbeitungsschritt um distanzmessspezifische Fehler korrigierten sowie auf die Horizontalstrecken  $D_\perp$  reduzierten Distanzwerte zwischen dem jeweiligen Pixel  $\mathbf{x}'$  und dem korrespondierenden Objektpunkt  $\mathbf{X}$  zur Verfügung (vgl. [Abschnitt 8.3](#) und [Kapitel 9](#)). Sie ermöglichen die Aufstellung einer weiteren Beobachtungsgleichung pro Pixel:

$$g_1^{D_\perp}(x', y') - e^{D_\perp}(x', y') = d_0 + d_1 \cdot g_2^{D_\perp}(x', y') \quad (14.4)$$

Die innerhalb eines diskreten Zeitschritts auf einen zu verfolgenden Bildpunkt wirkenden Distanzunterschiede werden in dem vorliegenden Ansatz linear modelliert. In der Beobachtungsgleichung (14.4) berücksichtigt der additive Korrekturterm  $d_0$  auftretende Translationen in Tiefenrichtung. Der multiplikative Faktor  $d_1$  skaliert den Entfernungswert zusätzlich.

Die geometrische Relation (14.1) soll für beide Beobachtungsarten einer **3D-Kamera** gleichermaßen gelten, nämlich sowohl für die Amplitudenwerte als auch für die Distanzmessungen. Es muss

folglich hinterfragt werden, inwieweit ein Zusammenhang zwischen den Affintransformationsparametern  $a_i$  und  $b_i$  mit  $i = [0, 3]$  und den neu eingeführten Distanzkorrekturtermen  $d_{0,1}$  besteht. In den folgenden Teilabschnitten wird zunächst analysiert, inwieweit der Affintransformationsparametersatz den aus dem Entfernungskanal extrahierten Oberflächenausschnitt beeinflusst, sodass Korrelationen mit  $d_0$  und  $d_1$  zu erwarten sind. Ist dies der Fall, wird sogleich der entsprechende Distanzkorrekturparameter als Funktion des mit ihm korrelierten Affintransformationsparameters ausgedrückt.<sup>8</sup>

#### Parametrisierung mit Translations- und Rotationskoeffizienten

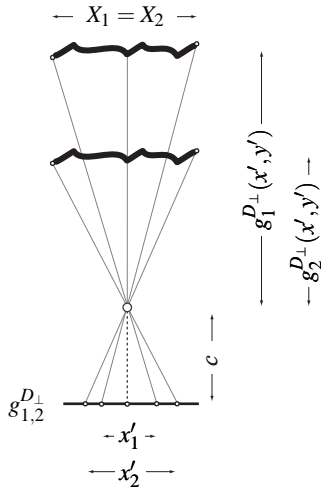
Reine Translationen des zu verfolgenden Objekts parallel zur Bildebene verursachen, ebenso wie ebene Rotationen, keine Variationen in den gemessenen Distanzwerten. In beiden Fällen werden die Distanzmessungen lediglich entsprechend ihrer transformierten sowie meist nicht ganzzahligen Bildposition interpoliert und der Inhalt des Suchfensters neu aufgestellt. Es besteht also kein Zusammenhang zwischen  $(a_0, a_2, b_0, b_1)$  und der Distanzverschiebung  $d_0$  bzw. -skalierung  $d_1$ .

#### Parametrisierung mit Maßstabskoeffizienten

Anders verhält es sich bei Maßstabsvariationen zwischen dem Muster- und Suchfenster, welche entweder durch Oberflächendeformationen oder durch Abstandsänderungen zwischen Objekt und Sensor hervorgerufen werden. Sind die Objektdeformationen zwischen zwei konsekutiven Bildern aufgrund der hohen zeitlichen Auflösung des Sensors und der zu untersuchenden kontinuierlichen menschlichen Bewegungen nicht signifikant, beeinflussen nur noch die Tiefenvariationen des Objekts selbst die affinen Maßstabskoeffizienten  $(a_1, b_2)$ . Diese Änderungen der Entfernung zwischen einem Musterfenster  $g_1^{D_\perp}$  und einem Suchfenster  $g_2^{D_\perp}$  werden in dem entwickelten Ansatz durch den *Offset*-Parameter  $d_0$  modelliert und in der folgenden Herleitung als Funktion der Maßstabsparameter ausgedrückt. Spätere Korrelationen bei der Lösung des funktionalen Modells sind dadurch vermeidbar.

Die parallel zur  $x'$ -Achse verlaufende Seitenlänge  $X_2$  eines in den Objektraum projizierten Suchbildausschnitts  $g_2^{D_\perp}$  entspricht bei einer reinen Verschiebung in Aufnahme­richtung der Seitenlänge  $X_1$  des Musterbildausschnitts  $g_1^{D_\perp}$ , erfährt jedoch im Bildraum eine perspektivbedingte Größenänderung von  $x'_1$  zu  $x'_2$  (**Abbildung 14.2**). Für die Bildmaßstabzahlen  $m_{1,2}$  gelten folglich die folgenden Verhältnisse zwischen den Seitenlängen im Bild- und Objektraum bzw. den gemessenen Objektstrecken  $g_{1,2}^{D_\perp}(x', y')$  und der Kamerakonstante  $c$ :

<sup>8</sup> Selbstverständlich wäre an dieser Stelle auch umgedreht die Formulierung der Affintransformationsparameter als Funktion des jeweiligen Distanzkorrekturterms möglich, jedoch aus Gründen der folgenden Softwareimplementierung weniger zweckmäßig.



$$m_i = \frac{x'_i}{X_i} = \frac{c}{g_i^{D_{\perp}}(x', y')} \quad (14.5)$$

mit  
 $i = [1, 2]$

**Abb. 14.2:** Bildmaßstabszahl: Zusammenhang zwischen konstanter Objektstrecke, variierender Aufnahmeentfernung und projizierter Strecke im Bildraum.

Da  $x'_1 \neq x'_2$  und  $X_1 = X_2$  folgt:

$$\frac{x'_2}{x'_1} = \frac{g_1^{D_{\perp}}(x', y')}{g_2^{D_{\perp}}(x', y')} \quad (14.6)$$

Tritt lediglich eine Objektverschiebung in Tiefenrichtung auf, vereinfacht sich das erweiterte affine Transformationsmodell (14.1) zu:

$$x'_2 = a_1 \cdot x'_1 \quad (14.7)$$

Da Objektdeformationen ausgeschlossen wurden, gelten die geführten Betrachtungen in gleicher Weise für die  $y'$ -Richtung und erlauben die Einführung eines in Zeilen- und Spaltenrichtung identischen Maßstabsparameters  $\lambda := \frac{1}{2}(a_1 + b_2)$ .

Aus den Gleichungen (14.6) und (14.7) ergibt sich der Zusammenhang zwischen dem affinen Maßstabsparameter  $\lambda$  und einer Translation in Tiefenrichtung zu:

$$\lambda = \frac{g_1^{D_{\perp}}(x', y')}{g_2^{D_{\perp}}(x', y')} \quad (14.8)$$

Aufgrund maßstabsinvarianter Nachbarschaftsdifferenzen innerhalb der Distanzmatrix gilt Gleichung (14.8) nur für das Centerpixel  $(x'_C, y'_C)$ :

$$g_1^{D_{\perp}}(x'_C, y'_C) = \lambda \cdot g_2^{D_{\perp}}(x'_C, y'_C) \quad (14.9)$$

Die transformierten Entfernungswerte der verbleibenden Nachbarschaft  $(x', y')$  ergeben sich zu:

$$g_1^{D_\perp}(x', y') = g_1^{D_\perp}(x'_C, y'_C) + \left[ g_2^{D_\perp}(x', y') - g_2^{D_\perp}(x'_C, y'_C) \right] \quad (14.10)$$

Durch die Substitution von  $g_1^{D_\perp}(x', y')$  aus Gleichung (14.10) mit dem Ausdruck (14.9) folgt sogleich die Distanzverschiebung  $d_0$  für jedes Pixel der Distanzmatrix als Funktion des Maßstabs  $\lambda$ :

$$d_0 = f(a_1, b_2) = g_1^{D_\perp}(x', y') - g_2^{D_\perp}(x', y') = g_2^{D_\perp}(x'_C, y'_C) \cdot (\lambda - 1) \quad (14.11)$$

Schließlich kann die für kleine Oberflächenelemente geltende Beobachtungsgleichung (14.4) neu aufgestellt werden:

$$g_1^{D_\perp}(x', y') - e^{D_\perp}(x', y') = g_2^{D_\perp}(x'_C, y'_C) \cdot \left( \frac{a_1 + b_2}{2} - 1 \right) + d_1 \cdot g_2^{D_\perp}(x', y') \quad (14.12)$$

#### Parametrisierung mit Neigungskoeffizienten

Eine Objekttranslation in Tiefenrichtung beeinflusst alle Pixel der Distanzmatrix in gleichem Maße und wird über die additive Distanzverschiebung  $d_0$  erfasst. Darüber hinaus kann es zu Objektneigungen, also Rotationseffekten in Tiefenrichtung kommen, die bei der Auswertung zweidimensional abgebildeter Daten über die Zusatzparameter  $(a_3, b_3)$  modelliert werden (Abbildung 14.1). Solche Effekte in bzw. gegen die Aufnahmerichtung wirken sich selbstverständlich auch auf die gemessenen Distanzwerte in Form von zeilen- bzw. spaltenabhängigen Maßstabsvariationen innerhalb der Distanzmatrix aus. Skaliert nun der oben eingeführte multiplikative Distanzkorrekturfaktor  $d_1$  die Entfernungswerte  $D_\perp$  zwischen den Bildausschnitten  $g_1^{D_\perp}$  und  $g_2^{D_\perp}$  in Abhängigkeit von ihrer jeweiligen Position  $(x', y')$  in beide Koordinatenrichtungen, kann ein analytischer Zusammenhang zwischen  $d_1$  und den Neigungskoeffizienten  $(a_3, b_3)$  hergestellt und, in gleicher Weise wie bei der Relation zwischen  $d_0$  und  $(a_1, b_2)$ , das Distanzkorrekturmodell (14.4) mit dem erweiterten Affintransformationsmodell (14.1) verknüpft werden. Im Folgenden wird  $d_1$  als Funktion von  $(a_3, b_3)$  und  $(x', y')$  hergeleitet.

Der Wertebereich für  $a_3$  bzw.  $b_3$  bei einer maximal abbildbaren Neigung um die  $x'$ - bzw.  $y'$ -Achse beträgt:

$$-\frac{2}{R} \leq a_3 \leq \frac{2}{R} \quad \text{und} \quad -\frac{2}{C} \leq b_3 \leq \frac{2}{C} \quad (14.13)$$

Für  $a_3 = 2/R$ , also eine maximalen Neigung um die  $x'$ -Achse mit  $y'_1 = y'_2 = y' = \text{const}$ , werden zunächst die Auswirkung auf die erste ( $y' = R/2$ ), die mittlere ( $y' = 0$ ) und die letzte ( $y' = -R/2$ ) Zeile der Distanzmatrix analysiert (Abbildung 14.3). Unter Vernachlässigung aller übrigen Effekte ( $a_0 = a_2 = b_0 = b_1 = 0$ ;  $a_1 = b_2 = 1$ ), gilt laut Transformationsgleichung (14.1):

$$x'_2 = x'_1 + a_3 \cdot x'_1 \cdot y'_1 \quad (14.14)$$

bzw. eingesetzt

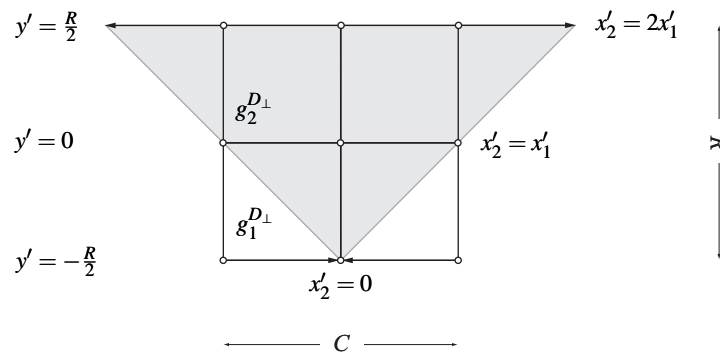
$$x'_2 = \begin{cases} 2x'_1 \\ x'_1 \\ 0 \end{cases} \quad \text{für} \quad y'_1 = \begin{cases} R/2 \\ 0 \\ -R/2 \end{cases} \quad (14.15)$$

Es wird deutlich, dass es maximal zu einer scheinbaren Verdoppelung der Pixelgröße in  $x'$  aufgrund der Neigung gegen die Aufnahme­richtung kommen kann. Da diese auftretenden Maßstabsänderungen real durch eine Kippung des Objekts zum Sensor verursacht werden, muss der Bildinhalt des in seiner Größe nicht zu verändernden Suchfensters  $g_2^{D\perp}$  entsprechend angepasst, in diesem Fall also halbiert werden. Die in der Neigungsachse gelagerte Zeile erfährt hingegen keine Änderungen; die zugehörigen Entfernungswerte bleiben somit ebenfalls unberührt. Die transformierten  $x'$ -Koordinaten der untersten Zeile fallen für diesen Extremfall bei  $x' = 0$  zusammen und erfahren aufgrund der Linearität der Transformation eine Vergrößerung ihrer gespeicherten Distanzwerte. Analytisch ausgedrückt bedeutet dies:

$$g_1^{D\perp}(x', y') = d_{1x'} \cdot g_2^{D\perp}(x', y') \quad (14.16)$$

mit

$$d_{1x'} = \begin{cases} 1/2 \\ 1 \\ 1,5 \end{cases} \quad \text{für} \quad y'_1 = \begin{cases} R/2 \\ 0 \\ -R/2 \end{cases}$$



**Abb. 14.3:** Maximal abbildbare Neigung um die  $x'$ -Achse.

In einem nächsten Schritt wird  $d_{1x'}$  neben  $y'$  zusätzlich als Funktion von  $a_3 = 2/R$  ausgedrückt:

$$d_{1x'} = \begin{cases} R/4 \cdot a_3 \\ R/2 \cdot a_3 \\ -R/4 \cdot a_3 \end{cases} = \begin{cases} 1/2 \\ 1 \\ 1,5 \end{cases} \quad \text{für} \quad y'_1 = \begin{cases} R/2 \\ 0 \\ -R/2 \end{cases} \quad (14.17)$$

Um die Relation (14.17) für beliebige Zeilen  $y'$  formulieren zu können, wird z. B. zwischen den

Funktionswerten  $d_{1x'} = f(y') = f(0)$  und  $f(R/2)$  linear interpoliert:

$$\begin{aligned} d_{1x'} = f(y') &= f(0) + \frac{f\left(\frac{R}{2}\right) - f(0)}{\frac{R}{2} - 0}(y' - 0) \\ &= 1 - \frac{y'}{2} \cdot a_3 \end{aligned} \quad (14.18)$$

Die gleichen Betrachtungen können für eine Kippung um die  $y'$ -Achse mit dem Neigungskoeffizienten  $b_3$  angestellt werden. Dies führt analog (und ohne Herleitung) zu:

$$d_{1y'} = f(x') = 1 - \frac{x'}{2} \cdot b_3 \quad (14.19)$$

Für den Gesamteinfluss des Distanzkorrekturfaktors folgt schließlich:

$$d_1 = f(\mathbf{x}', a_3, b_3) = \left(1 - \frac{y'}{2} \cdot a_3\right) \cdot \left(1 - \frac{x'}{2} \cdot b_3\right) \quad (14.20)$$

### Zusammenfassung

Es ergibt sich für die Beobachtungsgleichung (14.4) des Entfernungskanals die folgende allgemeine Form in Abhängigkeit der oben diskutierten und für die Anpassung des Distanzbildinhalts relevanten Affintransformationsparameter:

$$\begin{aligned} g_1^{D\perp}(x', y') - e^{D\perp}(x', y') &= d_0 + d_1 \cdot g_2^{D\perp}(x', y') \\ &= d_0 + d_{1x'} \cdot d_{1y'} \cdot g_2^{D\perp}(x', y') \\ &= g_2^{D\perp}(x'_C, y'_C) \cdot \left(\frac{a_1 + b_2}{2} - 1\right) + \left(1 - \frac{y'}{2} \cdot a_3\right) \cdot \left(1 - \frac{x'}{2} \cdot b_3\right) \cdot g_2^{D\perp}(x', y') \end{aligned} \quad (14.21)$$

Unter Bildung der Distanzgradienten  $g_{x'}^{D\perp}$  und  $g_{y'}^{D\perp}$  folgt nach einer Reihenentwicklung die lineari-



sierte Beobachtungsgleichung des Distanzkanals:

$$\begin{aligned}
g_1^{D\perp}(x',y') - e^{D\perp}(x',y') &= \left(g_2^{D\perp}\right)^0(x',y') + g_{x'}^{D\perp} \cdot da_0 \\
&+ g_{x'}^{D\perp} \cdot x' \cdot da_1 + \frac{1}{2} \cdot \left(g_2^{D\perp}\right)^0(x'_C, y'_C) \\
&+ g_{x'}^{D\perp} \cdot y' \cdot da_2 \\
&+ g_{x'}^{D\perp} \cdot x' \cdot y' \cdot da_3 + \frac{y'}{2} \cdot \left(g_2^{D\perp}\right)^0(x', y') \cdot \left(\frac{x'}{2} \cdot b_3 - 1\right) \\
&+ g_{y'}^{D\perp} \cdot db_0 \\
&+ g_{y'}^{D\perp} \cdot x' \cdot db_1 \\
&+ g_{y'}^{D\perp} \cdot y' \cdot db_2 + \frac{1}{2} \cdot \left(g_2^{D\perp}\right)^0(x'_C, y'_C) \\
&+ g_{y'}^{D\perp} \cdot x' \cdot y' \cdot db_3 + \frac{x'}{2} \cdot \left(g_2^{D\perp}\right)^0(x', y') \cdot \left(\frac{y'}{2} \cdot a_3 - 1\right)
\end{aligned} \tag{14.22}$$

mit

$$g_{x'}^{D\perp} = \partial \left(g_2^{D\perp}\right)^0(x', y') / \partial x', \quad g_{y'}^{D\perp} = \partial \left(g_2^{D\perp}\right)^0(x', y') / \partial y'$$

Die Bereitstellung von Näherungswerten für die Translations- und Rotationsparameter ( $a_0, a_2, b_0, b_1$ ) entspricht der Vorgehensweise des [Abschnitts 14.2.1](#). Die zusätzlich eingeführten Neigungsparameter ( $a_3, b_3$ ) werden mit Null initialisiert ( $d_1 = 1$ ). Die Translation  $d_0$  in Tiefenrichtung kann nach Gleichung [\(14.11\)](#) über

$$d_0^0 = g_1^{D\perp}(x'_C, y'_C) - \left(g_2^{D\perp}\right)^0(x'_C, y'_C) \tag{14.23}$$

bereits an eine bestmögliche Lösung approximiert werden und stellt nach den Gleichungen [\(14.18\)](#) bzw. [\(14.19\)](#) gleichzeitig Näherungswerte für die Maßstabsparameter bereit:

$$a_1^0 = b_2^0 = \frac{d_0^0}{\left(g_2^{D\perp}\right)^0(x'_C, y'_C)} + 1 \tag{14.24}$$

### 14.2.3 Anzahl der Beobachtungen und Unbekannten

Jedes Element der  $(R \times C)$  px großen Amplituden- und Distanzbildausschnitte liefert jeweils eine Beobachtung pro Kanal. Die Gesamtanzahl  $n$  setzt sich somit zusammen aus:

$$n = n_A + n_{D\perp} = 2 \cdot R \cdot C \tag{14.25}$$

Die unbekannt radiometrischen Parameter werden, aufgrund möglicher Korrelationen mit den

geometrischen Parametern, vor der Ausgleichung geschätzt. Somit verbleiben maximal  $u = 8$  unbekannte geometrische Korrekturterme  $(a_i, b_i)$  mit  $i = [0, 3]$  des erweiterten Affintransformationsparametersatzes.

#### 14.2.4 Koeffizientenmatrix

Die linearisierten Beobachtungsgleichungen (14.3) und (14.22) des Amplituden- und Distanzkanals seien als  $\Phi^A$  und  $\Phi^{D\perp}$  bezeichnet. Ihre Koeffizienten  $\left(\frac{\partial\Phi}{\partial\mathbf{X}}\right)^0$  werden zur Aufstellung der Koeffizientenmatrix  $\mathbf{A}$  herangezogen:

$$\mathbf{A}_{(n,u)} = \begin{bmatrix} \mathbf{A}_A & \mathbf{A}_{D\perp} \\ (n_A, u) & (n_{D\perp}, u) \end{bmatrix}^T \quad (14.26)$$

mit

$$a^A = F \left[ \left( \frac{\partial\Phi^A}{\partial a_i} \right)^0, \left( \frac{\partial\Phi^A}{\partial b_i} \right)^0 \right]$$

$$a^{D\perp} = F \left[ \left( \frac{\partial\Phi^{D\perp}}{\partial a_i} \right)^0, \left( \frac{\partial\Phi^{D\perp}}{\partial b_i} \right)^0 \right]$$

#### 14.2.5 Beobachtungsvektor

Der verkürzte Beobachtungsvektor  $\mathbf{l}$  beinhaltet die Differenzen zwischen den tatsächlichen Beobachtungen, also den Amplitudenwerten  $g_1^A$  bzw. den Distanzwerten  $g_1^{D\perp}$  an den Stellen  $(x'_1, y'_1)$ , und den Funktionswerten  $g_2^A$  bzw.  $g_2^{D\perp}$ , berechnet an den Stellen  $(x'_2, y'_2)$  mit den Näherungen des Parametersatzes  $(a_i, b_i)$  mit  $i = [0, 3]$ :

$$\mathbf{l}_{(n,1)} = \begin{bmatrix} \mathbf{l}_A & \mathbf{l}_{D\perp} \\ (n_A, 1) & (n_{D\perp}, 1) \end{bmatrix}^T \quad (14.27)$$

mit

$$l_{rc}^A = g_1^A(x'_{1c}, y'_{1r}) - g_2^A(x'_{2c}, y'_{2r})$$

$$l_{rc}^{D\perp} = g_1^{D\perp}(x'_{1c}, y'_{1r}) - d_1 \cdot g_2^{D\perp}(x'_{2c}, y'_{2r}) - d_0$$

### 14.3 Stochastisches Modell

Das stochastische Modell fasst die Genauigkeitsinformationen des funktionalen Modells zusammen (Abschnitt 4.4.2). Die Zuverlässigkeit und die Wertigkeit der heterogenen Amplituden- und Distanzbeobachtungen werden in Form der Varianzen  $\sigma_A^2$  bzw.  $\sigma_{D\perp}^2$  bereitgestellt und im gemeinsamen Ausgleichungsprozess des integrierten 2,5D-LST in der Kovarianzmatrix  $\Sigma_{ll}$  der Beobachtungen berücksichtig-

sichtigt:

$$\mathbf{\Sigma}_{ll} = \text{diag} \left( \begin{array}{c} \mathbf{\Sigma}_A \\ (n_A, n_A) \end{array}, \begin{array}{c} \mathbf{\Sigma}_{D_\perp} \\ (n_{D_\perp}, n_{D_\perp}) \end{array} \right) \quad (14.28)$$

mit

$$\begin{aligned} \sigma_{ll}^A &= s_A^2 \\ \sigma_{ll}^{D_\perp} &= s_{D_\perp}^2 \end{aligned}$$

Die a-posteriori Standardabweichung  $\hat{s}_{D_\perp}$  der ursprünglichen Distanzmessungen als ein Ergebnis der integrierten Bündelblockausgleichung fällt bei optimal veranschlagten Gewichten mit der a-priori Standardabweichung  $s_{D_\perp}$  zusammen ( $s_0 = \hat{s}_0$ ; Gleichung 4.31).  $\hat{s}_{D_\perp}$  wurde bereits optimal für die zu einer Beobachtungsart zusammengefassten Entfernungsmessungen im Rahmen der VKS einer integrierten Bündelblockausgleichung geschätzt (Tabelle 11.7) und kann folglich direkt im Modell (14.28) mit  $s_{D_\perp} = 10$  mm angesetzt werden. Für die Amplitudenbeobachtungen liegen zu diesem Zeitpunkt noch keine Informationen vor. In einer diese Arbeit begleitenden Genauigkeitsuntersuchung (M. Hempel 2007) wird, über Wiederholungsmessungen ermittelt, die mittlere Standardabweichung  $\bar{s}_A$  einer Amplitudenmessung mit ca. 40 GW angegeben.<sup>9</sup> Da  $\hat{s}_A = \bar{s}_A$ , ebenso wie  $\hat{s}_{D_\perp} = 1$  mm (Tabelle 11.7), nur noch zufällige Fehleranteile repräsentiert, und die Verhältnisse zwischen  $\hat{s}_A/\hat{s}_A$  und  $\hat{s}_{D_\perp}/\hat{s}_{D_\perp}$  gleich sind, approximiert  $\hat{s}_A = 400$  GW die a-priori Standardabweichung einer Amplitudenmessung.

Für eine auch im Fall des 2,5D-LST optimale Integration und vollständige Ausnutzung des Informationsgehalts beider Beobachtungsgruppen werden die zwei Gruppenvarianzen  $\hat{s}_A$  und  $\hat{s}_{D_\perp}$  automatisch im Zuge einer VKS geschätzt und daraus die Kovarianzmatrix  $\mathbf{\Sigma}_{ll}$  der Beobachtungen erzeugt. Dass aus statistischer Sicht eine Zusammenfassung gleichartiger Messwerte zu einer Beobachtungsgruppe sinnvoll ist, haben R. Hempel und Westfeld (2009) durch eine robuste Berechnung der Varianz-Kovarianzmatrix bewiesen (vgl. auch Abschnitt 11.6).

Als Ergebnis liegen über die Kovarianzmatrizen  $\mathbf{\Sigma}_{xx}$  der Unbekannten,  $\mathbf{\Sigma}_{\hat{y}}$  der ausgeglichenen Beobachtungen und  $\mathbf{\Sigma}_{vv}$  der Verbesserungen statistische Genauigkeitsmaße vor (Abschnitt 4.5.2). Sie werden zur Validierung des integrierten Tracking-Ansatzes im Abschnitt 14.5 verwendet.

## 14.4 Lösung der Ausgleichungsaufgabe

Im Folgenden wird der Ablauf des entwickelten 2,5D-LST Ansatzes kurz vorgestellt, der Datenfluss erläutert und auf Besonderheiten bei der Implementierung hingewiesen. Die Abbildung 14.4 fasst die Lösung der Ausgleichungsaufgabe in einer kompakten Darstellung der Programmstruktur zusätzlich zusammen.

<sup>9</sup> Die Amplitudendaten einer 3D-Kamera liegen in einer digitalen Auflösung von 16bit vor. Die maximal erfassbare Intensität eines Bildelements beträgt folglich 65.535 Grauwerte (GW).

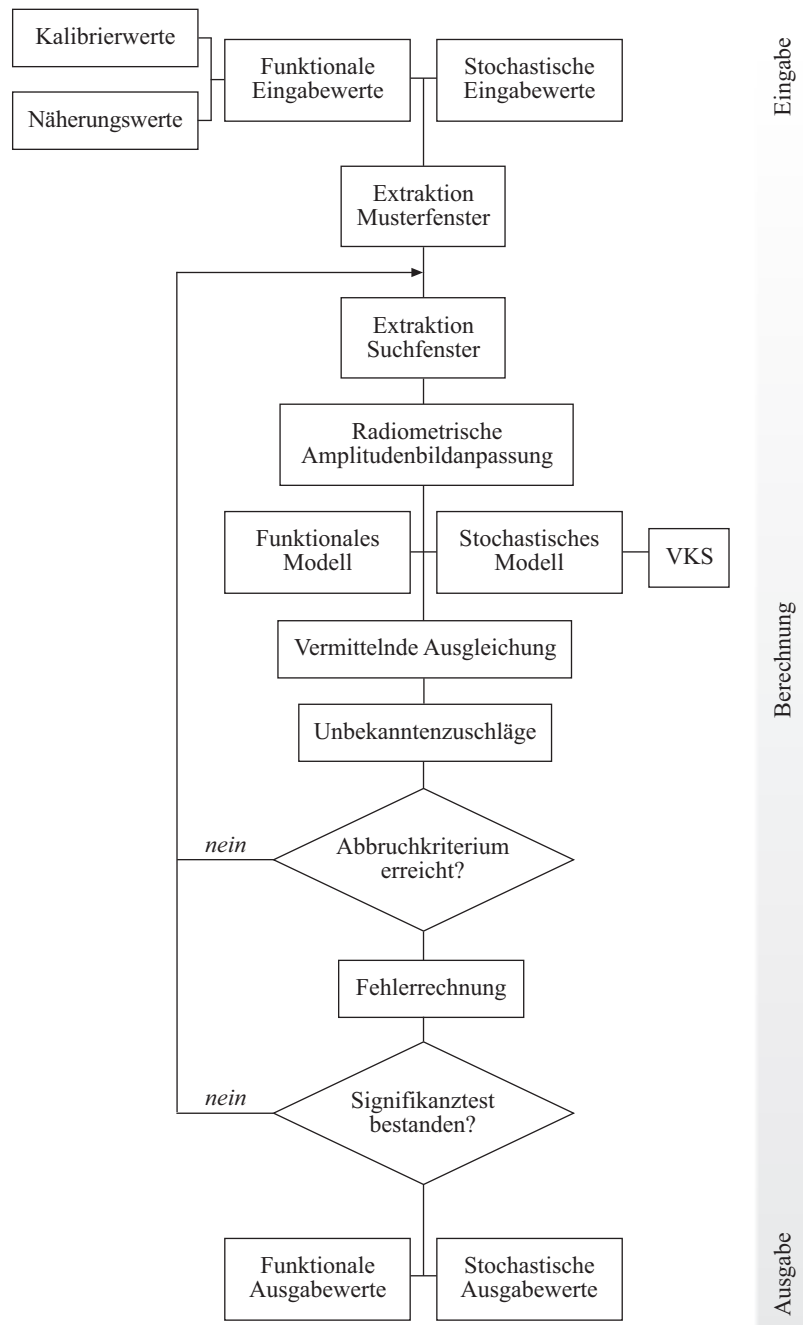


Abb. 14.4: Programmstruktur des integrierten 2,5D-LST.

#### 14.4.1 Eingabe

Die konsekutiven Amplituden- und Entfernungsbildpaare  $I_{1,2}^A$  und  $I_{1,2}^D$  entstammen einer 3D-Kamerabildsequenz. Sie werden direkt zur Aufstellung des funktionalen Modells herangezogen, wobei zunächst unter der Verwendung der in Kapitel 11 bestimmten Kalibrierparameter die Amplituden-

bilder um ihre Abbildungsfehler zu korrigieren und die Schrägdistanzen auf Horizontaldistanzen zu reduzieren sind (Abschnitt 8.3).

Für die im Laufe der MKQ zu lösenden linearisierten Beobachtungsgleichungen (14.3) und (14.22) sind Näherungswerte  $\mathbf{X}^0$  für alle unbekannt geometrischen Parameter notwendig, deren Bereitstellung im Abschnitt 14.2 bereits angesprochen und im Folgenden für die praktische Umsetzung konkretisiert wird:

- Die Translationsparameter  $(a_0, b_0)$  werden mit Null initialisiert. Es wird somit angenommen, dass es nur zu kleinen Bewegungen zwischen einem Zeitschritt kommt. Alternativen und in der Praxis verbreitete Vorgehensweisen im Falle einer zu geringen zeitlichen Auflösung des Sensors und/oder einer zu großen Objektbewegung wurden im Abschnitt 7.3.2 diskutiert und werden im Rahmen der Validierung des Punktverfolgungsalgorithmus nicht weiter betrachtet.
- Die Näherung für den in  $x'$  und  $y'$  einheitlichen Maßstabsparameter  $\lambda = \frac{1}{2}(a_1 + b_2)$  wird über die Gleichung (14.24) automatisch berechnet.
- Es wird von geringfügigen Rotations- und Scherungseffekten von einem Bild auf das zeitlich folgende Bild ausgegangen und die Parameter  $(a_2, b_1)$  entsprechend mit Null approximiert.
- Gleiches gilt für die zusätzlich zu modellierenden Neigungseffekte senkrecht zur Aufnahme-richtung. Bei kontinuierlichen (menschlichen) Bewegungen ist auch hier eine Initialisierung mit  $(a_3, b_3)$  gleich Null ausreichend.

Neben den funktionalen Größen sind die folgenden stochastischen Eingabewerte bereitzustellen:

- A-priori Standardabweichung der Gewichtseinheit  $s_0$
- A-priori Standardabweichungen der ursprünglichen Beobachtungen  $s_A$  und  $s_{D_\perp}$

#### 14.4.2 Extraktion der Muster- und Suchbilder

Als flächenbasiertes Verfahren bedient sich 2,5D-LST der um ein zuzuordnendes Bildelement  $\mathbf{x}'$  zentrierten Nachbarschaft. Für beide Kanäle werden die Musterfenster  $g_1^A$  und  $g_1^{D_\perp}$  einmalig gespeichert. Die zu transformierenden Suchfenster  $g_2^A$  und  $g_2^{D_\perp}$  müssen hingegen nach jeder Verbesserung des Unbekanntenvektors, also für jeden MKQ-Iterationsschritt  $k$ , neu extrahiert und bikubisch entsprechend der aktuellen und in der Regel nicht ganzzahligen Bildpositionen interpoliert werden.

#### 14.4.3 Radiometrische Anpassung

Vor der eigentlichen Hauptausgleichung werden im Rahmen einer weiteren MKQ der additive Helligkeitsparameter  $r_0$  und der multiplikative Kontrastkorrekturterm  $r_1$  berechnet. Die Normalgleichungs-

matrix  $\mathbf{N}$  und der Absolutgliedvektor  $\mathbf{n}$  lauten:

$$\begin{aligned}\mathbf{N} &= \mathbf{A}^T \mathbf{A} = \begin{bmatrix} R \cdot C & \sum^R \sum^C g_1^A(x'_c, y'_r) \\ \sum^R \sum^C g_1^A(x'_c, y'_r) & \sum^R \sum^C [g_1^A(x'_c, y'_r)]^2 \end{bmatrix} \\ \mathbf{n} &= \mathbf{A}^T \mathbf{l} = \begin{bmatrix} \sum^R \sum^C g_1^A(x'_c, y'_r) - g_2^A(x'_c, y'_r) \\ \sum^R \sum^C g_1^A(x'_c, y'_r) [g_1^A(x'_c, y'_r) - g_2^A(x'_c, y'_r)] \end{bmatrix}\end{aligned}\quad (14.29)$$

Entsprechend dem Formelapparat aus [Abschnitt 4.4.2](#) ergeben sich die gesuchten radiometrischen Korrekturparameter zu:

$$\hat{\mathbf{x}} = \mathbf{N}^{-1} \mathbf{n} = \begin{bmatrix} r_0 \\ r_1 - 1 \end{bmatrix}\quad (14.30)$$

Die Anpassung der Amplitudenwerte des Suchfensters  $g_2^A$  auf den Bildinhalt des Musterfensters  $g_1^A$  vollzieht sich entsprechend Gleichung (14.2):

$$\hat{g}_2^A(x', y') = \begin{cases} r_0 + r_1 \cdot g_2^A(x', y') & 0 \leq \hat{g}_2^A(x', y') \leq 2^{16} - 1 \\ 2^{16} - 1 & \text{für } \hat{g}_2^A(x', y') > 2^{16} - 1 \\ 0 & \hat{g}_2^A(x', y') < 0 \end{cases}\quad (14.31)$$

#### 14.4.4 Modellaufstellung

Die in den extrahierten Muster- und Suchfenstern gespeicherten Beobachtungen werden zur Aufstellung des funktionalen Modells entsprechend der [Abschnitte 14.2.4](#) und [14.2.5](#) verwendet. Die Amplitudenbildinhalte wurden bereits im vorherigen Teilabschnitt aufeinander angepasst. Die hierzu äquivalente Distanzbildanpassung mittels den neu eingeführten Distanzkorrekturtermen  $d_0$  und  $d_1$ , als Funktion der Maßstabs- und Neigungsparameter  $(a_2, b_2)$  bzw.  $(a_3, b_3)$ , geschieht bei der Modellbildung über die partiellen Ableitungen der Koeffizientenmatrix  $\mathbf{A}$  (Gleichungen [14.22](#) und [14.26](#)) sowie über die verbleibenden Distanzwertdifferenzen im verkürzten Lösungsvektor  $\mathbf{l}$  (Gleichung [14.27](#)). Lediglich der zu transformierende Distanzwert  $g_2^{D_\perp}(x'_c, y'_c)$  des Centerpixels  $\mathbf{x}'_c$  wird direkt für die Aufstellung beider Matrizen benötigt und folglich zuvor entsprechend der aktuellen Translationsparameter  $(a_0, b_0)$  – alle weiteren geometrischen Parameter haben im Mittelpunkt des Fensters keine Auswirkungen – bikubisch an der nach Gleichung (14.1) berechneten Position  $(x'_{2c}, y'_{2c})$  interpoliert. Die angepassten Distanzwerte aller anderen Elemente des Suchfensters ergeben sich aufgrund maßstabsinvarianter Nachbarschaftsdifferenzen indirekt entsprechend Gleichung (14.21). Die Matrizen  $\mathbf{A}$  und  $\mathbf{l}$  sind für jeden Iterationsschritt  $k$  neu aufzustellen.

Zusätzlich werden die Kovarianzmatrix  $\Sigma_{ll}$  der Beobachtungen bzw. die Gewichtsmatrix  $\mathbf{P}$  über die im Rahmen einer [VKS \(Abschnitt 4.6\)](#) nach ([Förstner 1979](#); [Koch 2004](#)) geschätzten Varianzen  $s_A^2$  und  $s_{D_\perp}^2$  zu jedem Iterationsschritt  $k$  gebildet.

#### 14.4.5 MKQ

Die Koeffizientenmatrix  $\mathbf{A}$ , der Beobachtungsvektor  $\mathbf{l}$  und die Gewichtsmatrix  $\mathbf{P}$  stellen alle notwendigen funktionalen und stochastischen Informationen bereit, um die Ausgleichungsaufgabe (4.20) im **Gauß-Markov-Modell (GMM)** zu lösen. Die geschätzten Parameterzuschläge  $\hat{\mathbf{x}}$  verbessern den aktuellen Parametervektor  $\mathbf{X}^k$  in jeder Iteration  $k$ .

#### 14.4.6 Abbruchkriterium

Es wird so lange iteriert, bis

$$1 - 1e^{-5} < \frac{(\mathbf{v}^T \mathbf{P} \mathbf{v})_k}{(\mathbf{v}^T \mathbf{P} \mathbf{v})_{k-1}} < 1 + e^{-5} \quad (14.32)$$

oder bis eine zuvor definierte maximale Anzahl an Iterationen erreicht ist. Eine Oszillation des Lösungsvektors führt ebenfalls zum Abbruch.

#### 14.4.7 Fehlerrechnung

Die Fehlerrechnung gestaltet sich nach dem in **Abschnitt 4.5** gegebenen Berechnungsapparat. Aus den Kovarianzmatrizen der Beobachtungen, Unbekannten und Verbesserungen werden folgende Qualitätsmaße abgeleitet und zur späteren Beurteilung der Punktzuordnung herangezogen (vgl. **Abschnitt 14.5**):

- A-posteriori Standardabweichung  $\hat{\sigma}_0$  der Gewichtseinheit
- A-posteriori Standardabweichungen  $\hat{\sigma}_l$  bzw.  $\hat{\sigma}_f$  der ursprünglichen bzw. ausgeglichenen Beobachtungen
- A-posteriori Standardabweichungen  $\hat{\sigma}_{\hat{\mathbf{x}}}$  der geschätzten unbekanntem Modellparameter
- A-posteriori Standardabweichungen  $\hat{\sigma}_{f(\hat{\mathbf{x}})}$  von Funktionen der geschätzten unbekanntem Modellparameter ( $d_0, d_1$ )
- A-posteriori Standardabweichungen  $\hat{\sigma}_v$  der Verbesserungen
- Koeffizienten  $\rho_{ij}$  zur Analyse möglicher Korrelationen zwischen den geschätzten Parametern
- Die Elemente der Matrix  $\mathbf{D}$  als Maße für die Bestimmbarkeit eines Parameters

#### 14.4.8 Signifikanztest

Zur Beurteilung der Signifikanz eines Parameters wird das Quantil  $t_{p,f}$  der t-Verteilung für einen zweiseitigen Vertrauensbereich über die Wahrscheinlichkeit  $p = 0,95$  sowie die Anzahl an Freiheitsgraden  $f$  berechnet (**Abramowitz & Stegun 1964**). Ist der Quotient  $\hat{x}_i/\hat{\sigma}_{\hat{x}_i}$  kleiner als  $t_{p,f}$ , wird der Parameter eliminiert und die Ausgleichung erneut durchlaufen. Vergleiche hierzu auch **Abschnitt 4.5.6**.



### 14.4.9 Ausgabe

Schließlich werden die folgenden Ergebnisse übergeben:

- Geschätzte Modellparameter  $\hat{x}_i$  mit ihren a-posteriori Standardabweichungen  $\hat{s}_{\hat{x}_i}$
- Residuen  $v_i$  der Beobachtungen
- Genauigkeits- und Zuverlässigkeitsmaße, berechnet im Zuge der oben genannten VKS und Fehlerrechnung

## 14.5 Validierung des integrierten Least Squares Tracking

Das entwickelte integrierte 3D-Kamerabildzuordnungsverfahren 2,5D-LST wird anhand verschiedener simulierter und realer Datensätze umfassend getestet (Abschnitt 14.5.1). Hauptsächliches Interesse galt der Validierung der aufgestellten funktionalen Zusammenhänge. Dabei spielen die folgenden beiden Fragestellungen eine besondere Rolle:

1. Bewirkt die Hinzunahme des Entfernungskanals als redundanter Informationsträger einen Gewinn für den Prozess der Bildzuordnung (Abschnitt 14.5.2)?
2. Ist eine zusätzliche Parametrisierung durch Neigungskoeffizienten sinnvoll (Abschnitt 14.5.3)?

Weiterhin werden die verbleibenden Residuen auf Restsystematiken untersucht (Abschnitt 14.5.4) und das Konvergenzverhalten des Zuordnungsprozesses dokumentiert (Abschnitt 14.5.5). Abschließend erfolgt die Analyse des stochastischen Modells (Abschnitt 14.5.6).

Die Entwicklung eines Verfahrens zur Bestimmung realer Referenzdaten zur Einschätzung der äußeren Genauigkeit sowie anwendungsspezifische Genauigkeitsuntersuchungen nach bspw. den Beurteilungskriterien von [Barnard und Fischler \(1982\)](#); vgl. auch [Abschnitt 15.3](#)) werden nicht durchgeführt und bleiben zukünftigen Arbeiten vorbehalten.

### 14.5.1 Experimentalkonfiguration

#### Datensätze

Simulierte Datensätze mit zufällig überlagertem Gaußschen Rauschen enthalten keine systematischen und nicht kontrollierbaren Messunsicherheiten aufgrund von bspw. Unstetigkeiten der Objekt-oberfläche, Hintergrundstrahlung oder Mehrwegeeffekten. Mehrere wie in [Abbildung 14.5](#) dargestellte Amplituden- und Distanzbilder werden synthetisch erzeugt und definiert in eine zweite Epoche transformiert. Die auf dieser Datengrundlage gerechneten Kleinste-Quadrate-Anpassungen schätzen die gesuchten 2,5D-LST-Affintransformationsparameter korrekt; lediglich das hinzugefügte Messrauschen verbleibt in den a-posteriori Standardabweichungen  $\hat{s}_{\hat{A},\hat{D}}$  der ausgeglichenen Beobachtungen. Sie bestätigen somit die prinzipielle Eignung des aufgestellten funktionalen und stochastischen Modells. Die Ergebnisse werden mit Verweis auf ([Westfeld & Hempel 2008](#)) an dieser Stelle nicht weiter dokumentiert.

Als Zweites werden zwei reale Datensätze von verschiedenen Standpunkten generiert ([Abbildung 14.6](#)). Das dabei aufgenommene statische Referenzobjekt besitzt im Amplituden- und im Entfernungskanal sowohl kontrastarme als auch kontrastreiche Regionen, was die Grundlage in der Beurteilung des Zugewinns komplementärer Informationen bildet. Weiterhin ermöglichen die aus unterschiedlichen Entfernungen und aus verschiedenen Richtungen erzeugten Aufnahmen Aussagen über die Parametrisierung des geometrischen Modells. Die Muster- und Suchbildaufnahmen des ersten Datensatzes sind signifikant zueinander um die  $x'$ - und  $y'$ -Achse geneigt und weisen mit  $g_1^{D\perp}(x'_C, y'_C) = 1,3\text{ m}$  und  $g_2^{D\perp}(x'_C, y'_C) = 1,5\text{ m}$  einen geringen Entfernungsunterschied auf. Das Musterbild des zweiten Datensatzes ist aus 1,1 m aufgenommen worden; das ebenfalls um beide Koordinatenachsen geneigte Suchbild aus 2,3 m.

Abschließend wird **2,5D-LST** an einer 145 Bilder umfassenden Sequenz einer im Mittel ca. 1,5 m entfernten Winkbewegung einer Hand validiert ([Abbildungen 14.7](#) und [14.8](#)). Diese weist eine natürliche und für die im Rahmen der Arbeit fokussierte Anwendung der menschlichen Bewegungsanalyse typische Kontrastverteilung in beiden Kanälen auf.

Alle akquirierten 3D-Kameraaufnahmen werden unter der Verwendung der in [Kapitel 11](#) ermittelten Kalibrierparameter entsprechend ihrer Abbildungsfehler korrigiert. Die Bilddaten des Entfernungskanals werden zusätzlich um die Distanzmessfehler korrigiert und nach den Vorgaben des in [Abschnitt 14.2.2](#) aufgestellten funktionalen Modells auf Horizontaldistanzen reduziert.

### Einstellungen

Die Modulationsfrequenz  $f$  der 3D-Kamera PMD[vision]<sup>®</sup> CamCube 2.0 beträgt 20 MHz. Eine mit 2,0 ms angesetzte Integrationszeit  $\Delta t$  minimiert Unschärfe bei der Aufnahme bewegter Objekte, gewährleistet aber gleichzeitig eine für Aufnahmebereiche bis maximal ca. 3,5 m ausreichende Infrarotausleuchtung. Die zusätzlich zur Unterdrückung von Bewegungsartefakten hardwareseitig implementierte **SMB-Funktion** (*Suppression of Motion Blur*) bleibt vorerst deaktiviert. Eine Mittelbildung bei der Aufnahme der statischen Szene findet nicht statt.

**2,5D-LST** ist anfänglich stets voll parametrisiert mit  $(a_i, b_i)$ , wobei  $i = [0, 3]$ ;  $a_1 = b_2$ , und wird entsprechend dem nach jedem Zuordnungsschritt durchgeführten Signifikanztest angepasst ([Abschnitt 14.4.8](#)). Die Größe der zu transformierenden Bildausschnitte liegt bei  $25^2$  px. Somit stehen zur Aufstellung des hochredundanten **LST-Systems**  $n = 1.250$  Beobachtungen zur Verfügung (Gleichung [14.25](#)). Alle weiteren Einstellungen wurden bereits in [Abschnitt 14.4](#) beschrieben.

### 14.5.2 Verwendung komplementärer Entfernungsinformationen

Als flächenbasiertes differentiell arbeitendes Zuordnungsverfahren benötigt das entwickelte **2,5D-LST** Gradienten. Fehlt es den Kanälen an ausreichend Kontrast in möglichst beiden Koordinatenrichtungen, soll also bspw. ein ebenes und im **nahen Infrarot (NIR)** einheitlich reflektierendes Oberflächensegment zugeordnet werden, ist das Korrespondenzproblem nicht eindeutig oder gar nicht lösbar ([Abschnitt 7.3](#)). Der folgende Abschnitt analysiert, inwieweit sich das Konvergenzverhalten

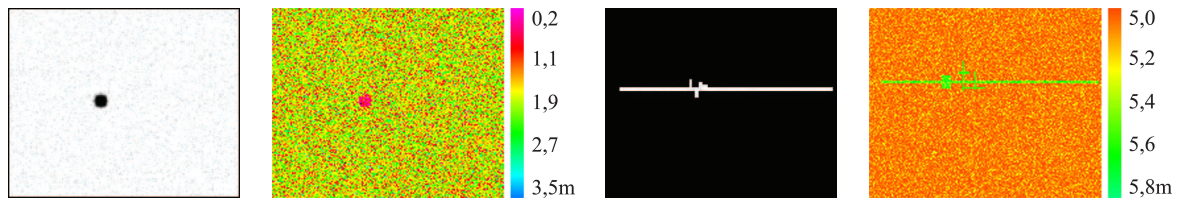


Abb. 14.5: Zwei Beispiele eines synthetisch erzeugten 3D-Kameradatensatzes.

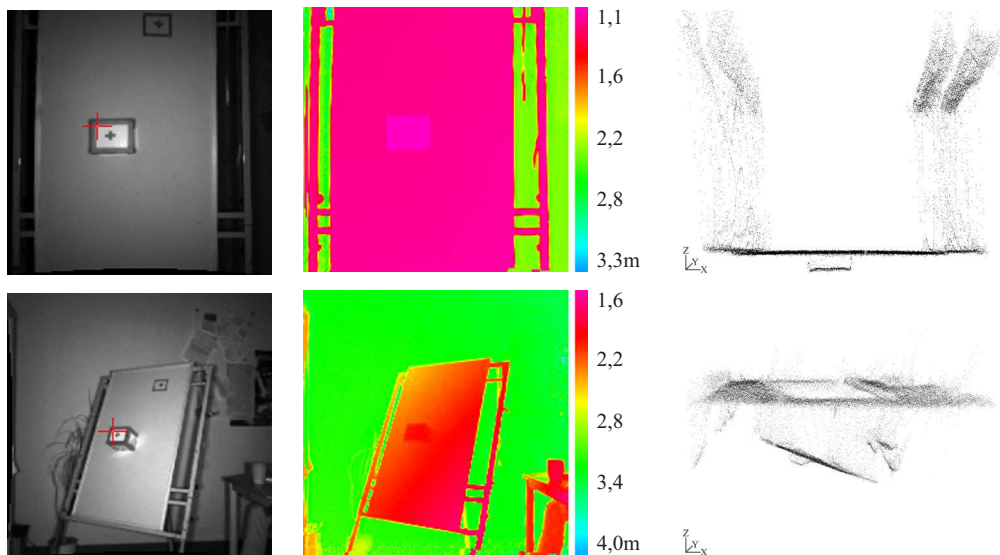


Abb. 14.6: 3D-Kamerabildaufnahmen und 3D-Punktwolkendarstellungen eines statischen Referenzobjekts von unterschiedlichen Standpunkten (Auswertereien i bis vi). Der zugeordnete Objektpunkt der Auswertung  $v$  ist rot markiert.

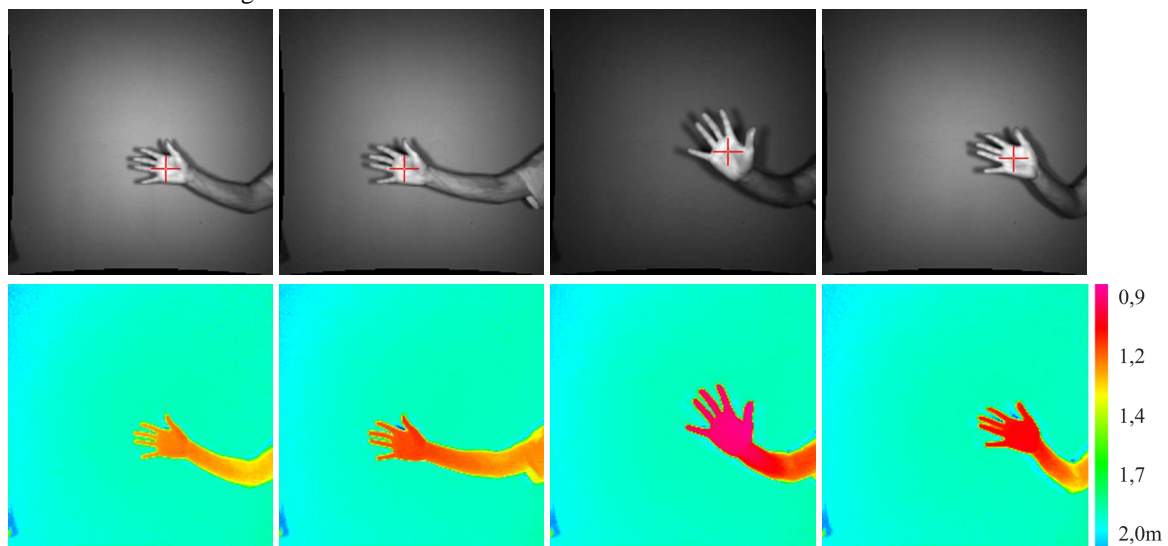
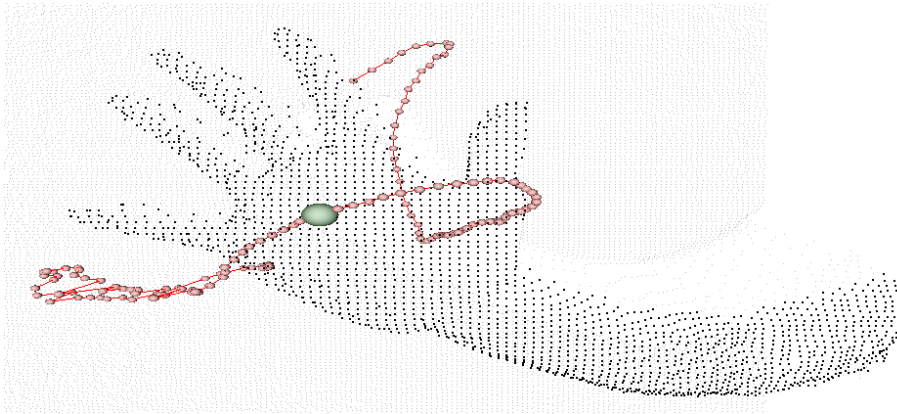


Abb. 14.7: 3D-Kamerabildsequenz einer sich bewegenden Hand zu den Zeitpunkten 1, 30, 70 und 145. Das rote Fadenkreuz markiert den in den Auswertereien vii und viii mit 2,5D-LST verfolgten Objektpunkt. Die gesamte Trajektorie ist in der folgenden Abbildung 14.8 dargestellt.



**Abb. 14.8:** Punktwolke zur 70. Epoche einer 145 Bilder umfassenden 3D-Kamerasequenz, überlagert mit der in der Auswertung viii gerechneten vollständigen Trajektorie des verfolgten Handpunkts.

und das damit einhergehende Genauigkeitsniveau des Bildzuordnungsalgorithmus unter der Verwendung komplementärer Beobachtungen verhält. Hierfür werden zwei Datensätze eines statischen Referenzobjekts erstellt und ausgewertet, welche vor allem in dem neu eingeführten Entfernungskanal unterschiedliche Kontrastverhältnisse aufweisen (Abschnitt 14.5.1; Abbildungen 14.6 und 14.7; Tabelle 14.1, Spalte *Objekt*).

Die in der Tabelle 14.1 dokumentierten 2,5D-LST-Auswertereien i und ii bilden ein texturiertes ebenes Oberflächensegment aufeinander ab, welches zwischen den Epochen um die  $x', y'$ -Bildachsen geneigt wird. Im Vergleich zu einer einkanaligen Amplitudenbildzuordnung kann durch die Hinzunahme des Entfernungskanals keine Steigerung der Translations-, Rotations- und Maßstabsparametergenauigkeiten erreicht werden. Vielmehr bewirken die signalschwachen Entfernungsinformationen eine Verschlechterung um ca. 25% (vgl. Abschnitt 4.5.3); in den a-posteriori Standardabweichungen  $\hat{s}_{a_0, b_0}$  der Verschiebungen bspw. von  $1/28$  px auf  $1/20$  px. Einzig die neu eingeführten und über den Distanzkorrekturfaktor  $d_1$  mit den Entfernungsmessungen verknüpften Neigungskoeffizienten  $(a_3, b_3)$  können um fast eine Größenordnung genauer geschätzt werden, was sich letztlich in ähnlichen a-posteriori Standardabweichungen  $\hat{s}_A$  und  $\hat{s}_{\hat{\lambda}}$  widerspiegelt.

Die Zuordnung kleiner texturierter aber nicht gänzlich ebener Oberflächensegmente widerspricht dem Prinzip einer 2D-Affintransformationsabbildung. Die durch Höhenunterschiede hervorgerufenen perspektiven Verzerrungen können nur teilweise aber keinesfalls geometrisch adäquat über die Rotations- und Maßstabsparameter modelliert werden (Abschnitt 7.6.1). Die Genauigkeitsangabe der Auswertereihe iv belegen dies quantitativ: Zwar konvergiert die mit Amplitudeninformationen angesetzte Kleinste-Quadrate-Anpassung nach 27 Iterationen, jedoch nicht an der korrekten Position. Abhilfe schafft die Verwendung komplementärer Entfernungsinformationen. Die Gradienten im Entfernungsbild stützen den Zuordnungsprozess bzw. ermöglichen erst eine zuverlässige Schätzung der Transformationsparameter unter einem hinreichenden Genauigkeitspotential (Auswertereihe v; Abbildung 14.14).

Die Kontrastverhältnisse der Auswertereien vii und viii sind ebenfalls ausgewogen, jedoch im

Entfernungskanal aufgrund der Verfolgung einer Hand im Unterschied zu iv bis vi weniger stark ausgeprägt. Folglich verbessern sich vor allem die Genauigkeiten der mit den Distanzkorrekturtermen  $d_0$  und  $d_1$  verknüpften Maßstabs- sowie Neigungsparameterpaare  $(a_1, b_2)$  und  $(a_2, b_1)$  um ca. 15%.

**Tab. 14.1:** Die Tabelle zeigt Genauigkeitsangaben für die **2,5D-LST**-Zuordnung verschieden ausgerichteter Objekte (Spalte *Adjust*). Dabei sind die folgenden a-posteriori Standardabweichungen dokumentiert:  $\hat{s}_{a_i, b_i}$  der unbekanntem Affintransformationsparameter  $(a_i, b_i)$  mit  $i = [0, 3]$ ,  $\hat{s}_{d_{0,1}}$  der Funktionen  $d_0 = f(a_1, b_2)$  und  $d_1 = f(a_3, b_3)$  sowie  $\hat{s}_l$  und  $\hat{s}_r$  der ursprünglichen und ausgeglichenen Beobachtungen. Es finden Amplitudeninformationen oder Amplituden- und Entfernungsmessungen Verwendung (A oder AD; Spalte *Kanal*). Die letzte Spalte *It.* gibt die zur Konvergenz des Lösungsvektors notwendige Iterationsanzahl an.

Für die ohne **VKS** durchgeführten Auswertereien i, iv und vii ist die a-posteriori Standardabweichung  $\hat{s}_0$  der Gewichtseinheit gleich  $\hat{s}_A$ ; andernfalls ist  $s_0 = \hat{s}_0 = 1,00$ .

Die Angaben der Auswertereihe vii resultieren aus der arithmetischen Mittelung der Genauigkeitsangaben aller 145 Bilder der Sequenz.

ID	Objekt	Adjust	Kanal	$\hat{s}_{a_0, b_0}$ [px]	$\hat{s}_{a_1, b_2}$	$\hat{s}_{a_2, b_1}$	$\hat{s}_{a_3, b_3}$	$\hat{s}_{d_0}$ [% $D_{\perp}$ ]	$\hat{s}_{d_1}$	$\hat{s}_A$ [GW]	$\hat{s}_{\hat{A}}$	$\hat{s}_{D_{\perp}}$ [mm]	$\hat{s}_{\hat{D}_{\perp}}$	It.
i	eben	Neigung	A	$3,59e^{-2}$	$3,81e^{-3}$	$7,52e^{-3}$	$8,20e^{-4}$	-	-	227	18	-	-	13
ii			AD	$4,88e^{-2}$	$5,13e^{-3}$	$9,33e^{-3}$	$1,04e^{-4}$	0,24	$8,71e^{-4}$	356	21	14	1,13	11
iii			AD	$5,30e^{-2}$	$5,50e^{-3}$	$1,14e^{-2}$	fix	0,27	fix	356	23	28	1,30	9
iv	3D	Neigung Entfernung	A	0,27	$2,10e^{-2}$	$6,54e^{-2}$	$6,50e^{-3}$	fix	fix	1509	119	fix	fix	27
v			AD	$5,53e^{-2}$	$5,50e^{-3}$	$7,51e^{-3}$	$8,84e^{-5}$	0,26	$7,28e^{-4}$	577	26	17	1,14	21
vi			AD	$7,56e^{-2}$	$6,85e^{-3}$	$1,00e^{-2}$	fix	0,32	fix	592	32	36	1,81	24
vii	Hand		A	$8,65e^{-3}$	$1,05e^{-3}$	$1,31e^{-3}$	$1,41e^{-4}$	fix	fix	57	4	-	-	7
viii			AD	$9,06e^{-3}$	$8,95e^{-4}$	$1,07e^{-3}$	$1,23e^{-4}$	$7,23e^{-2}$	$9,18e^{-4}$	72	4	20	1,00	15

### 14.5.3 Parametrisierung

Der folgende Abschnitt analysiert die **2,5D-LST**-Parametrisierung. Das Hauptaugenmerk liegt dabei auf der Distanzverschiebung  $d_0$  als Funktion der affinen Maßstabsparameter sowie auf dem Einfluss der Distanzskalierung  $d_1$  als Funktion der neu eingeführten Neigungskoeffizienten. Zu diesem Zweck wird das Referenzobjekt der statischen Datensätze signifikant verschoben und geneigt (Tabelle 14.1, Spalte *Adjust*) sowie die Auswertung um die nicht die Neigungskoeffizienten  $(a_3, b_3)$  mitschätzenden Reihen iii und vi ergänzt. Die ursprünglichen **2D**-Affintransformationsparameter  $(a_0, b_0)$  zur Erfassung von Verschiebungen und  $(a_2, b_1)$  zur Modellierung von Rotationseffekten werden ebenso wie die auch beim konventionellen **2D-LSM** auftretenden Maßstabskoeffizienten  $(a_1, b_2)$  zunächst stets mitgeschätzt und ggf. entsprechend ihrer Signifikanz später ausgeschlossen.

Die Translationsparameter  $(a_0, b_0)$  sind zur Berechnung von Verschiebungsvektoren notwendig und somit photogrammetrisch von großem Interesse. Ihre a-posteriori Standardabweichung  $\hat{s}_{a_0, b_0}$



kann im quadratischen Mittel mit ca.  $1/20$  px angegeben werden (Tabelle 14.1). Dieser Wert liegt im Vergleich zu einem 2D-LSM im unteren Bereich der zu erwartenden Genauigkeit bei der Zuordnung kontrastreicher Zielmarken (Abschnitt 7.6.1). Angesichts der geringen räumlichen Auflösung einer 3D-Kamera und dem in beiden Kanälen ausgeprägt vorhandenen Signalrauschen (Kapitel 8) ist seine Größenordnung aber realistisch.

Der mit den Maßstabsparametern  $(a_1, b_2)$  in funktionalen Zusammenhang gebrachte *Offset*-Parameter  $d_0$  modelliert Distanzvariationen in Aufnahme­richtung und ergänzt so den lateralen Verschiebungsvektor  $\Delta \mathbf{x}'$  um die Tiefenkomponente. Die Auswertereien ii und v ordnen Oberflächensegmente in den Abständen 1,3 m bis 2,3 m einander zu. Die a-posteriori Standardabweichung  $\hat{s}_{d_0}$ <sup>10</sup> der additiven Distanz­korrektur beträgt im Mittel ca. 0,25 % der zwischen dem Centerpixel  $\mathbf{x}'_C$  und dem korrespondierenden Objektpunkt  $\mathbf{X}$  gemessenen Horizontal­distanz  $D_{\perp}^C$ . Bei bspw. einer relativen Verschiebung von 272 mm in 1,6 m Entfernung ist  $\hat{s}_{d_0} = 3,8$  mm.

Die aufgestellten Parameter  $(a_3, b_3)$  zur Modellierung von Neigungen um beide Bildkoordinatenachsen sowie der daraus gerechnete Distanz­korrekturfaktor  $d_1$  können mit einer Wahrscheinlichkeit von 95 % signifikant geschätzt werden. Für z. B. die obere linke Ecke  $(-13; 13)$  des Suchbildausschnitts resultiert der für die Auswertereihe ii geschätzte Neigungskoeffizient  $a_3 = -2,63e^{-3}$  nach Gleichung (14.18) in einem Distanz­korrekturfaktor  $d_{1x'} = 1 + 1,534e^{-2}$ , für  $b_3 = -3,8e^{-3}$  in  $d_{1y'} = 0,977$ . Zusammengefasst bewirkt der Distanz­korrekturfaktor  $d_1 = d_{1x'} \cdot d_{1y'} = 0,9917$  nach (14.20) in 1,5 m Entfernung eine Verschiebung von ca.  $-13$  mm. Die a-posteriori Standardabweichung  $\hat{s}_{d_1}$ <sup>11</sup> beträgt in diesem Fall  $8,71e^{-4}$ . Die Auswertereien iii und vi werden ohne  $(a_3, b_3)$  parametrisiert (Tabelle 14.1). Vor allem die Genauigkeiten  $\hat{s}_i$  der verbleibenden Transformationsparameter verringern sich dadurch um ca. 20 %. Ebenso dokumentieren die im Vergleich zu ii und v um teils bis zu 50 % schlechter ausfallenden a-posteriori Standardabweichungen  $\hat{s}_{A, D_{\perp}}$  und  $\hat{s}_{\hat{A}, \hat{D}_{\perp}}$  der ursprünglichen und ausgeglichenen Beobachtungen ein weniger adäquat aufgestelltes Transformationsmodell. Es wird ersichtlich, dass insbesondere durch die Integration der Entfernungsbeobachtungen in die mit einem additiven und multiplikativen Distanz­korrekturterm parametrisierte lineare Beobachtungsgleichung (14.4) der funktionale Zusammenhang wirklichkeitsgetreuer modelliert werden kann.

Die a-posteriori Standardabweichungen der in der Auswertung viii geschätzten unbekannt­en Transformationsparameter (Winkbewegung einer Hand) fallen im Vergleich zu den Auswertungen der statischen Datensätze deutlich besser aus (Tabelle 14.1). Zu erklären ist dies mit einer kontinuierlich verlaufenden Bewegung und den damit einhergehenden geringen radiometrischen und geometrischen Änderungen zwischen zwei konsekutiven Bildern. Die Tabelle 14.2 zeigt darüber hinaus die Prozentangaben der signifikant bestimmten Transformationsparameter der Auswertereihe viii. Bei der Erfassung der Winkbewegung werden die Parameter  $(a_i, b_i)$  mit  $i = [0, 2]$  einer standardmäßig

<sup>10</sup>  $\hat{s}_{d_0}$  ist definiert als a-posteriori Standardabweichung der Funktion (14.11) der Unbekannten  $a_1$  und  $b_2$ ; vgl. auch Gleichung (4.34).

<sup>11</sup>  $\hat{s}_{d_1}$  ist definiert als a-posteriori Standardabweichung der Funktion (14.20) der Unbekannten  $a_2$  und  $b_1$  für den Maximalfall  $x' = -R/2$  und  $y' = C/2$ ; vgl. auch Gleichung (4.34).

mit jeweils zwei Translationen, Rotationen und Maßstäben parametrisierten 2D-Affintransformation in durchschnittlich mehr als 90% der gerechneten Ausgleichungen signifikant geschätzt. Die Prozentzahl für die neu eingeführten Parameter  $(a_3, b_3)$  zur Berücksichtigung möglicher Neigungen um die Bildachsen liegt bei deutlich über 50% und belegt abschließend den Nutzen einer erweiterten Parametrisierung.

**Tab. 14.2:** Prozentangaben der signifikant bestimmten Transformationsparameter in der 145 Epochen umfassenden Bildsequenz der Auswertereihe viii.

Parameter	$a_0$	$b_0$	$a_1, b_2$	$a_2$	$b_1$	$a_3$	$b_3$
%	96	97	87	90	83	55	61

#### 14.5.4 Residuenanalyse

Die Analyse des Residuenvektors  $\mathbf{v}$  gibt Aufschluss über die Richtigkeit des aufgestellten funktionalen Zusammenhangs und zeigt an, ob die verbleibenden Verbesserungen normalverteilt um ihren Erwartungswert  $\mu = 0$  streuen oder (noch zu modellierende) Restsystematiken im Datenmaterial vorhanden sind.

In der [Abbildung 14.9](#) sind exemplarisch am bewegungsanalytischen Beispiel der Auswertung viii die Häufigkeiten der Residuen der Amplituden- und Distanzbeobachtungen für einen Zuordnungsschritt dargestellt. Die Residuen  $v_A$  der Amplitudenmessungen streuen im Mittel mit  $\hat{s}_{v_A} = 56 \text{ GW}$  normalverteilt um den Erwartungswert  $\mu = -2,4 \text{ GW}$ . Die verbleibenden Restklaffungen  $v_{D_\perp}$  der Distanzmessungen weisen eine a-posteriori Standardabweichung  $\hat{s}_{v_{D_\perp}}$  von 11,74 mm auf. Sie streuen normalverteilt ohne erkennbare systematische Effekte um  $\mu = 2,50 \text{ mm}$ .

Die [Abbildung 14.11](#) zeigt ebenfalls die Häufigkeiten der Residuen der Amplituden- und Distanzbeobachtungen am Beispiel der Auswertung v des statischen Datensatzes. In diesem Fall werden bewusst die Verteilungsfunktionen einer räumlichen Zuordnung zweier aus stark unterschiedlichen Entfernungen und Richtungen aufgenommener Oberflächenausschnitte gewählt. Aufgrund der u. a. von der Sensorposition, der Entfernung und der Objektoberflächenbeschaffenheit abhängigen aktiven 3D-Kamerabeleuchtung ([Abschnitte 8.1.2](#) und [9.7](#)) bereitet die Anpassung des Amplitudenkanals unter solch extrem variierenden Beleuchtungsverhältnissen Probleme. Die Residuen  $v_A$  der Amplitudenmessungen streuen im Mittel mit  $\hat{s}_{v_A} = 573 \text{ GW}$  zwar normalverteilt, jedoch aufgrund verbleibender Restsystematiken verschoben um den Erwartungswert  $\mu = 156 \text{ GW}$ . Die [Abbildung 14.12](#) verdeutlicht zusätzlich die Erweiterungsfähigkeit der bis dato lediglich linear aufgestellten radiometrischen NIR-Amplitudenbildanpassung anhand von Differenzdarstellungen zwischen dem Musterfenster und den beiden Suchbildausschnitten vor bzw. nach der radiometrischen Anpassung. Bei einer zeitlich hochaufgelösten *Bild-zu-Bild*-Zuordnung (Auswertung viii; [Abbildung 14.10](#)) wirken sich die Effekte einer aktiven Beleuchtung angesichts der Kontinuität eines (menschlichen) Bewegungsablaufes weniger kritisch aus. Dennoch sollte die in Gleichung (14.2) unterstellte lineare radiometrische Relation zwischen Muster- und Suchbildausschnitt näher analysiert und ggf. angepasst werden. Im

Gegensatz dazu weisen die verbleibenden Restklaffungen  $v_{D_{\perp}}$  der Distanzmessungen keine erkennbaren systematischen Effekte auf. Sie streuen mit einer a-posteriori Standardabweichung  $\hat{s}_{v_{D_{\perp}}}$  von 17,70 mm normalverteilt um  $\mu = -1,90$  mm.

#### 14.5.5 Konvergenzverhalten

Die nichtlinearen Beobachtungsgleichungen (14.2) und (14.4) stellen beim 2,5D-LST den funktionalen Zusammenhang zwischen den Grau- und Entfernungswerten der Bildausschnitte  $g_{1,2}^{A,D_{\perp}}$  her (Abschnitt 14.2). Deren Linearisierung erfolgt an der Stelle der Näherungswerte  $\mathbf{X}^0$ , welche im Zuordnungsprozess iterativ durch das Anbringen der geschätzten Zuschläge  $\hat{\mathbf{x}}$  zum Modellparametersatz  $\mathbf{X}^k$  nach jedem Durchlauf  $k$  verbessert werden. Der Lösungsvektor  $\mathbf{X}$  konvergiert dabei in der Regel nach wenigen Durchläufen.

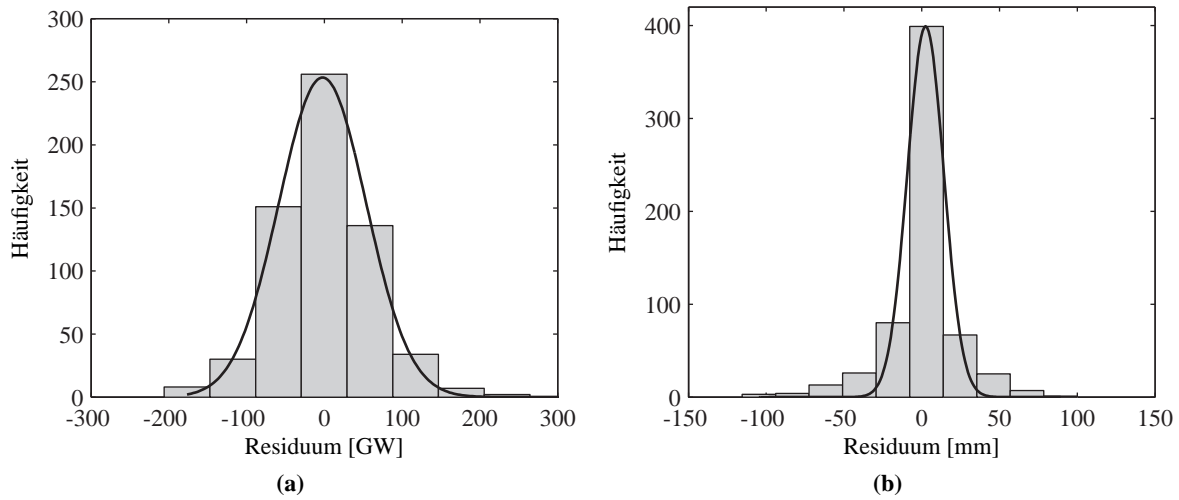
Die Abbildung 14.13 zeigt am Beispiel der Auswertung  $v$  des statischen Datensatzes (Tabelle 14.1) das Konvergenzverhalten aller eingeführten Affintransformationsparameter  $a_i$  und  $b_i$  mit  $i = [0, 3]$ . Die Abbildungsreihen 14.14 stellen die zugehörige geometrische Transformation des ursprünglich quadratischen Suchfensters sowie die entsprechend interpolierten Grau- und Entfernungswerte der jeweiligen Bildausschnitte dar.

Der Iterationsabbruch erfolgt nach dem 21. Durchlauf an der bestmöglichen Position unter der Minimierung der Summe der Quadrate der Grau- und Entfernungswertunterschiede. Die Änderung des Betrags der Verbesserungsquadratsumme (4.10) ist dabei entsprechend den in Abschnitt 14.4.6 definierten Kriterien kleiner  $1e^{-5}$ . Im Vergleich zu einer konventionellen Kleinste-Quadrate-Zuordnung konvergiert 2,5D-LST in diesem Beispiel etwas langsamer. Dies kann zum einen mit den zu jeder Iteration geschätzten Varianzkomponenten begründet werden. Zum anderen treten signifikante Änderungen der Transformationsparametern zwischen den zwei Aufnahmen unterschiedlicher Standpunkte auf, die lediglich mit den in Abschnitt 14.4.1 festgelegten Standardwerten initialisiert werden. Im Sinne einer vernünftigen Rechenzeit kann für praktische Anwendungen auf eine zu jeder Iteration durchgeführten VKS verzichtet werden. Bei der Bild-zu-Bild-Auswertung dynamischer Datensätze verbessert sich das Konvergenzverhalten ebenfalls, da die Transformationsparameter der vorangegangenen Epoche als gute Näherungswerte genutzt werden können. Bei der Auswertung des vorliegenden dynamischen Datensatzes werden inklusive VKS im Mittel 15 Durchläufe benötigt.

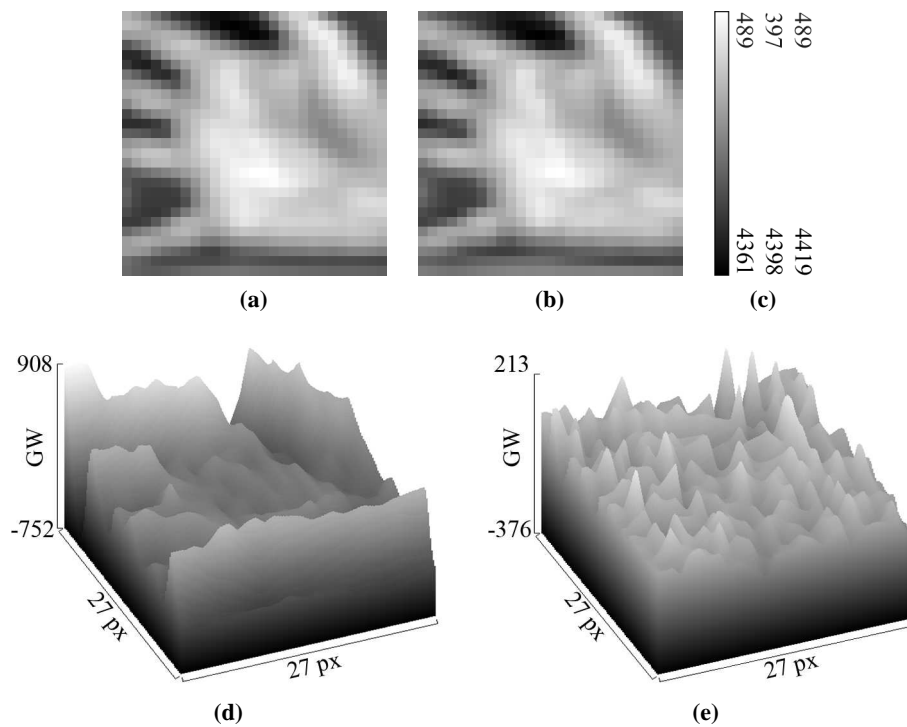
#### 14.5.6 Fehler der Beobachtungsgruppen

Die in Tabelle 14.1 gelisteten a-posteriori Standardabweichungen der Beobachtungen wurden für vergleichende Analysen in den Abschnitten 14.5.2 und 14.5.2 bereits quantifiziert. Die a-posteriori Standardabweichungen  $\hat{s}_{A,D_{\perp}}$  der ursprünglichen Beobachtungen sind die Ergebnisse einer zu jeder Iteration durchgeführten VKS (Abschnitt 14.3). Da die Messgenauigkeit der 3D-Kamerabeobachtungen primär von der Rückstrahlungsenergie der ausgesendeten NIR-Photonen abhängt (SNR), aber auch vom Reflexionsgrad der Objektoberfläche oder zusätzlich auftretender Mehrwegeeffekte, variieren deren Größenordnungen je nach Experimentalkonfiguration (vgl. auch Kapitel 8 und 9).

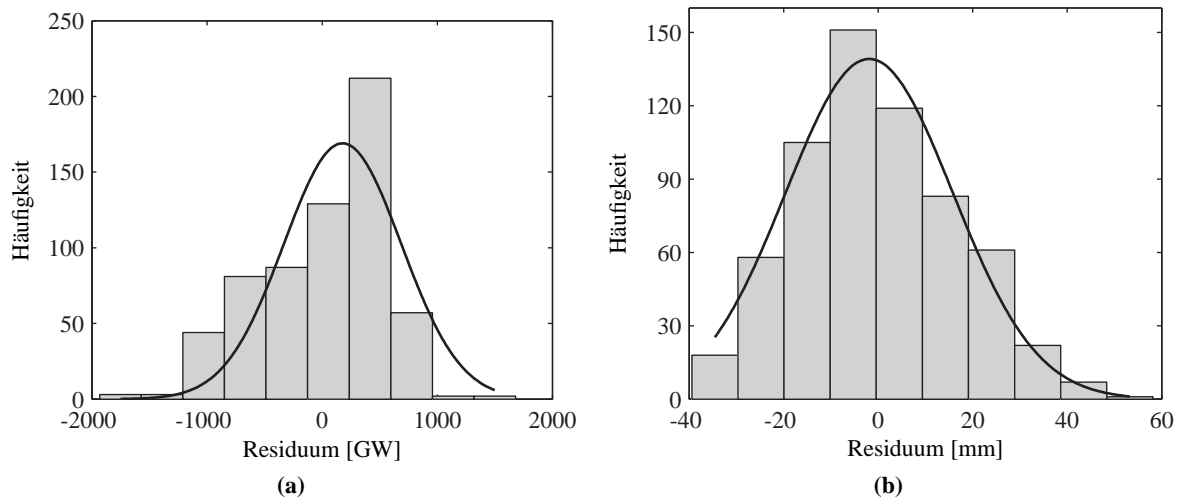




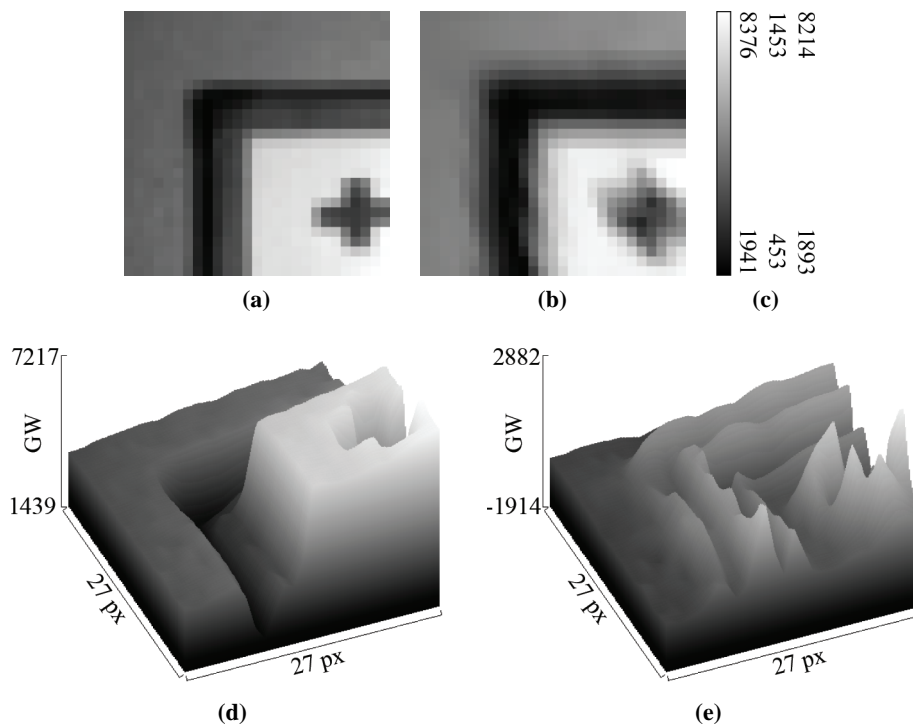
**Abb. 14.9:** Residuenverteilung, I: (a) Die Residuen  $v_A$  der Amplitudenmessungen der Auswertung viii streuen im Mittel mit  $\hat{s}_{v_A} = 56$  GW normalverteilt um den Erwartungswert  $\mu = -2,4$  GW. (b) Die verbleibenden Restklaffungen  $v_{D_{\perp}}$  der Distanzmessungen weisen eine a-posteriori Standardabweichung  $\hat{s}_{v_{D_{\perp}}}$  von 11,74 mm auf. Sie streuen normalverteilt ohne erkennbare systematische Effekte um  $\mu = 2,50$  mm.



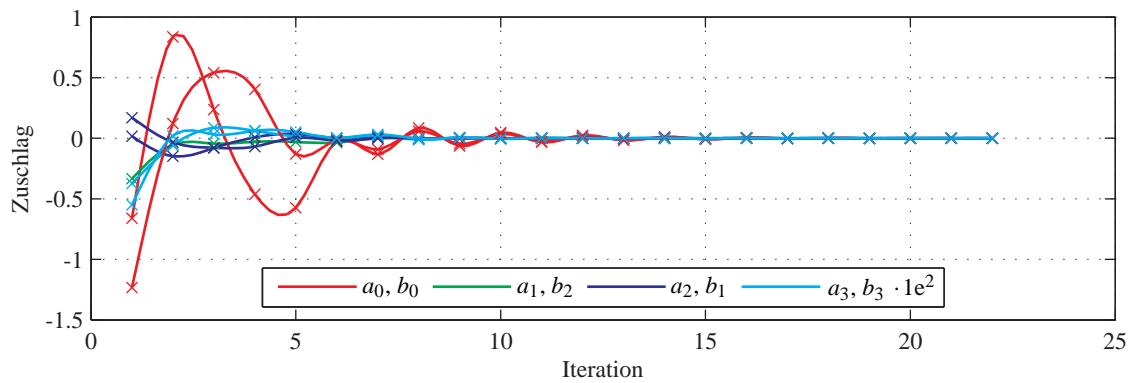
**Abb. 14.10:** Residuendarstellung, I: (a) und (b) zeigen den Muster- und bestmöglich transformierten Suchbildausschnitt des Amplitudenkanals der Auswertung viii. In (c) ist der zugehörige Wertebereich für das Musterbild (links) sowie für das Suchbild vor (Mitte) bzw. nach (rechts) einer linearen radiometrischen Transformation dargestellt. Die Plots (d) und (e) zeigen die Differenzen zwischen dem Musterfenster und den beiden Suchbildausschnitten vor bzw. nach der radiometrischen Anpassung.



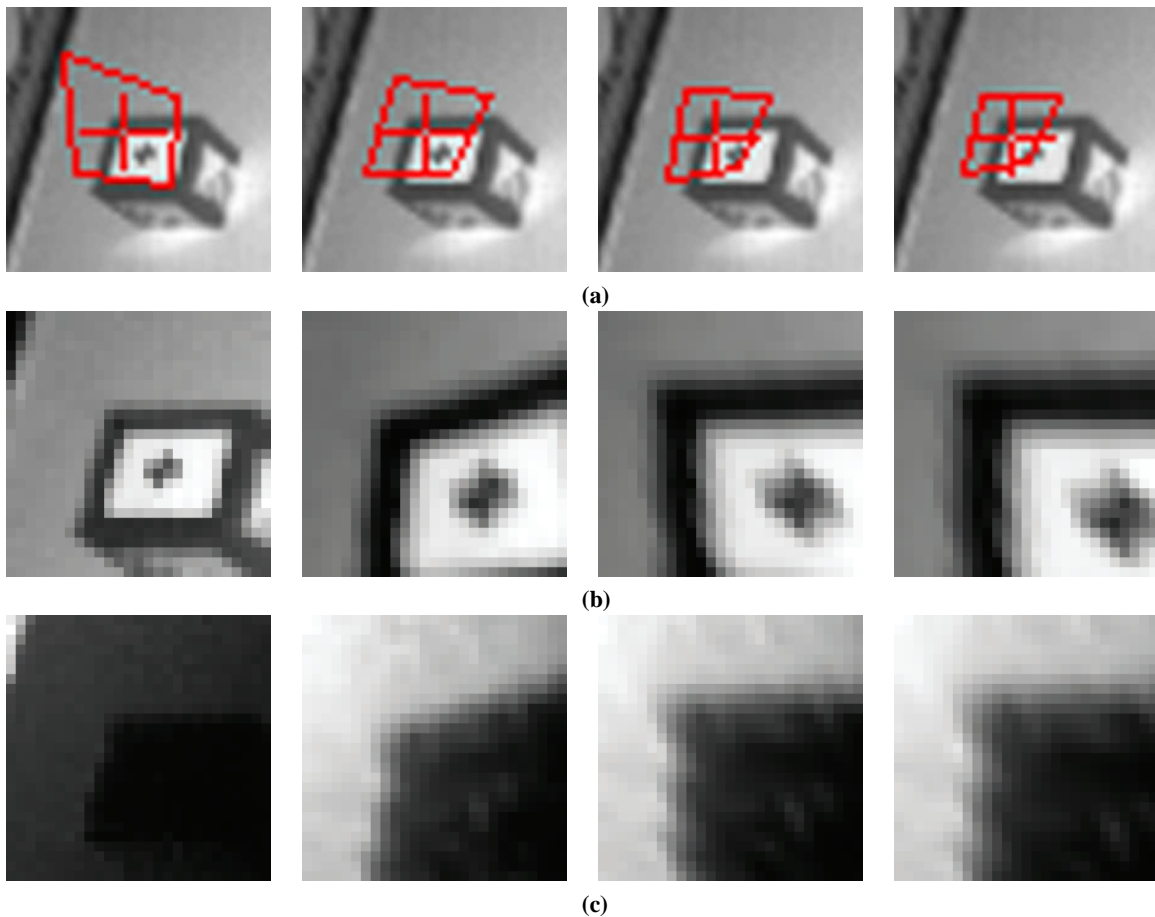
**Abb. 14.11:** Residuenverteilung, II: (a) Die Residuen  $v_A$  der Amplitudenmessungen der Auswertung  $v$  streuen im Mittel mit  $\hat{s}_{v_A} = 573$  GW normalverteilt um den aufgrund verbleibender Restsystematiken verschobenen Erwartungswert  $\mu = 156$  GW. (b) Die verbleibenden Restklaffungen  $v_{D\perp}$  der Distanzmessungen weisen eine a-posteriori Standardabweichung  $\hat{s}_{v_{D\perp}}$  von 17,70 mm auf. Sie streuen normalverteilt ohne erkennbare systematische Effekte um  $\mu_{\perp} = -1,90$  mm.



**Abb. 14.12:** Residuendarstellung, II: (a) und (b) zeigen den Muster- und bestmöglich transformierten Suchbildausschnitt des Amplitudenkanals der Auswertung  $v$ . In (c) ist der zugehörige Wertebereich für das Musterbild (links) sowie für das Suchbild vor (Mitte) bzw. nach (rechts) einer linearen radiometrischen Transformation dargestellt. Die Plots (d) und (e) zeigen die Differenzen zwischen dem Musterfenster und den beiden Suchbildausschnitten vor bzw. nach der radiometrischen Anpassung.



**Abb. 14.13:** Plot des Konvergenzverhaltens für ein voll parametrisiertes 2,5D-LST: Bereits nach wenigen Iterationen konvergiert der Zuordnungsprozess an der bestmöglichen Position. Die Änderung des Betrags der Verbesserungsquadratsumme (4.10) ist dabei kleiner  $1e^{-5}$ . Die Zuschlagsverläufe der Neigungsparameter ( $a_3, b_3$ ) sind mit  $1e^2$  überhöht dargestellt.



**Abb. 14.14:** Iterative Anpassung der Suchbildausschnitte: Die Abbildungsreihe (a) zeigt die geometrische Transformation des ursprünglich quadratischen Suchfensters, (b) und (c) visualisieren die entsprechend interpolierten Grau- und Entfernungswerte. Dargestellt ist der 1., 2., 5. und letzte (21.) Durchlauf. Das zugehörige Konvergenzverhalten der einzelnen Parameter ist in [Abbildung 14.13](#) dargestellt.

Für eine räumliche Bildzuordnung kann die Genauigkeit einer Amplitudenmessung im Mittel mit ca.  $\hat{s}_A = 500 \text{ GW}$  angegeben werden (Auswertereien iii und vi), was die a-priori angesetzte Standardabweichung  $s_A$  bestätigt (Abschnitt 14.3). In einer 8bit-Auflösung entspricht dies ca. 2GW. Die a-posteriori Standardabweichung  $\hat{s}_{\hat{A}}$  der ausgeglichenen Beobachtungen liegt bei im Mittel ca. 30GW (0,1 GW bezüglich 8 bit). Die a-posteriori Distanzmessgenauigkeiten belaufen sich für die ursprünglichen Beobachtungen auf ca. 30,0 mm.  $\hat{s}_{D_{\perp}}$  fällt im Vergleich zur angesetzten a-priori Standardabweichung  $s_{D_{\perp}}$  als ein Ergebnis der durchgeführten Kalibrierung um den Faktor 3 höher aus. Dies ist, im Gegensatz zur Messung und Modellierung weißer Styroporkugeln, mit der Zuordnung teils komplexer sowie aus unterschiedlichen Richtungen verschieden rückstrahlender Oberflächensegmente zu begründen. Mit  $\hat{s}_{\hat{D}_{\perp}} = 1,5 \text{ mm}$  bestätigt die a-posteriori Standardabweichung der ausgeglichenen Beobachtungen letztlich die Richtigkeit des aufgestellten funktionalen Zusammenhangs und entspricht in ihrer Größenordnung dem Ergebnis der integrierten Bündelblockausgleichung (Tabelle 11.7).

Erwartungsgemäß fallen die Genauigkeitsangaben für die zeitlichen Zuordnungsaufgaben zwischen jeweils zwei lediglich geringfügig variierenden konsekutiven Bildern deutlich besser aus, da der Einfluss der verbliebenen nicht oder nur schwer kalibrierbaren Messfehler gleich ist (Tabelle 14.1).

## 14.6 Diskussion

Der vorgestellte Algorithmus **2,5D-LST** ordnet mit einer **3D**-Kamera erfasste Oberflächensegmente automatisch spatial und/oder temporal zu. Zur Ausnutzung des gesamten Informationshaushalts wurden die zwei Beobachtungsarten des Amplituden- und Entfernungskanals geschlossen in einem integrierten geometrischen und stochastischen Modell formuliert. Als Ergebnisse liegen die geschätzten Transformationsparameter mit ihren statistischen Genauigkeitsangaben zur Beurteilung der Bewegungsschätzung vor. Von photogrammetrischem Interesse sind hier besonders die subpixel- bzw. submillimetergenau bestimmten Verschiebungsvektoren, welche für alle drei Koordinatenrichtungen vorliegen. Die als **Gauß-Helmert-Modell (GHM)** aufgestellte Zuordnungsmethode ist adaptiv; der vorgeschlagene Parametersatz kann flexibel erweitert oder um Zusatzwissen ergänzt werden. Ein auf **2,5D-LST** basierendes Verfahren kann als noninvasiv bezeichnet werden (Kapitel 15), da die entwickelte Methode zwar prinzipbedingt Kontrast in mindestens einem Messkanal benötigt, aber keine künstliche Signalisierung voraussetzt.

Zur Validierung des Potentials des entwickelten **2,5D-LST** wurden verschiedene Experimente durchgeführt. Zusammengefasst ist zu konstatieren, dass die zu den Amplitudenmessungen einer **3D**-Kamera komplementären Entfernungsinformationen genau dann als ein Zugewinn für die Ausgleichungsaufgabe gewertet werden können, wenn das Grauwertsignal (temporär) flach verläuft. Gleiches gilt für den umgekehrten Fall, also für die Stützung des Entfernungskanals mit geringen Distanzunterschieden durch kontrastreiche Amplitudeninformationen. Darüber hinaus verbessern die mit den Maßstabs- und neu eingeführten Neigungsparametern in funktionalen Zusammenhang gebrachten Distanzkorrekturterme die geometrische Transformation; vor allem, wenn die durch die

2D-Affintransformation geforderte Ebenheitsbedingung kleiner Bildausschnitte nicht erfüllt ist. Bei ausgeprägten Gradientenverläufen in beiden Kanälen werden zwar die Genauigkeiten der einzelnen Transformationsparameter nicht signifikant verbessert, jedoch die Robustheit und Zuverlässigkeit durch die Verwendung aller zur Verfügung stehenden Informationen erhöht. In diesem Zusammenhang ist die Implementierung eines Gradiententests vor der eigentlichen Ausgleichung sinnvoll, welcher z. B. von der mittleren quadratischen Abweichung der Grau- bzw. Entfernungswerte eines Bildausschnitts abhängt. Liegt die mittlere quadratische Abweichung des Bildinhalts unter einem zuvor definierten Schwellenwert, wird der zugehörige Kanal ausgeschlossen.

Die zur Schätzung optimaler Beobachtungsgewichte um eine VKS implementierte Erweiterung des stochastischen Modells ist sinnvoll, weil dadurch sowohl Beobachtungsarten unterschiedlicher Genauigkeiten als auch die von unterschiedlichen Experimentalanordnungen abhängenden Genauigkeitsniveaus adäquat berücksichtigt werden. Im praktischen Einsatz ist die Schätzung der Varianzkomponenten zu jeder Iteration aufgrund des erhöhten Rechenaufwands nicht zweckmäßig. Für eine gegebene Aufnahmekonfiguration sollten sie vielmehr einmalig bestimmt, für weitere Experimente als sehr gute Näherungen ansetzt und wenn notwendig, höchstens nach der Konvergenz des Lösungsvektors im Zuge einer zusätzlichen Iteration mit aktiver VKS erneut geschätzt werden. Dabei ist auch der Einfluss der Varianzkomponenten auf den Parametervektor zu untersuchen.

Die Analyse der inneren Genauigkeitsangaben, der verbleibenden Residuen, des Konvergenzverhaltens sowie der Beobachtungsfehler bestätigen zum einen die genaue und zuverlässige Arbeitsweise des entwickelten Zuordnungsansatzes, zum anderen die Richtigkeit des aufgestellten funktionalen Zusammenhangs. Einzig die Eignung des unterstellten linearen radiometrischen Modells zur Anpassung des Amplitudenkanals muss geprüft und ggf. um weitere bspw. die Sensorposition modellierende radiale Terme erweitert werden.

Im Mittel kann ein in einem Meter Entfernung aufgenommener Objektpunkt  $\mathbf{X}(X, Y, Z)$  mit einer inneren Genauigkeit von  $\hat{s}_{X,Y} = 0,15\text{ mm}$  und  $\hat{s}_Z = 1,0\text{ mm}$  zugeordnet werden. Für eine abschließende Beurteilung der Methode fehlen jedoch Aussagen über die äußere Genauigkeit. Im Rahmen zukünftiger Arbeiten sind hierfür Referenzdaten zu erzeugen. Diese können aus mit Sensoren übergeordneter Genauigkeit zu verschiedenen Epochen eingemessenen und in 3D-Kameradaten signalstark abgebildeten Zielmarken bestehen. Denkbar sind bspw. kleine dreidimensionale Referenzkörper nach dem Vorbild der in Kapitel 11 zur integrierten Kalibrierung verwendeten weißen Styroporkugeln. Ebenfalls wünschenswert ist ein Vergleich mit konkurrierenden *Tracking*-Algorithmen wie bspw. dem *Optical Range Flow* (Abschnitt 13.3.1).



## Anwendungen der nonverbalen Interaktionsforschung

Die Photogrammetrie hat sich in den vergangenen Jahren u. a. auch in den Bereichen Biomechanik, Rehabilitations- und Sportmedizin oder Filmanimation neue Anwendungsgebiete erschlossen (Kapitel 7). Kommerziell eingesetzte Bewegungsanalysesysteme bestehen aus bis zu 32 Kameras. Diese arbeiten invasiv auf der Basis der Verfolgung von durch retroreflektierende Zielmarken signalisierten Körperpunkten (Abschnitt 12.3). Eine 3D-Kamera bietet hier als monosensorielles Messsystem in Verbindung mit berührungslos arbeitenden Bildzuordnungsverfahren entscheidende Vorteile (Kapitel 7, 8 und 13).

Das nonverbale Verhalten reguliert im wesentlichen Maße implizit menschliche Interaktionen und besitzt einen großen Stellenwert in der menschlichen Verhaltensforschung (Patterson 1983). Als eine Dimension des nonverbalen Verhaltens gelten quantitativ zu erfassende Bewegungsabläufe. Hier liegt ein wichtiges praktisches Anwendungsgebiet des entwickelten flächenbasierten Zuordnungsverfahrens *2,5D Least Squares Tracking (LST)*. Zur Dokumentation des Verfahrenspotentials mit großer Relevanz in den Erziehungswissenschaften (Abschnitt 15.1) oder der Psychologie (Abschnitt 15.2) dient die hochaufgelöste Erfassung von Bewegungsabläufen in der Analyse nonverbaler Kommunikation in Interaktionsverläufen von Personen.

### 15.1 Konfliktregulation bei befreundeten Kindespaaren

#### 15.1.1 Motivation

Ein erziehungswissenschaftliches Untersuchungsgebiet beschäftigt sich mit der Frage, wie sich Freunde beeinflussen und voneinander lernen. In diesem Zusammenhang wird bspw. die Konfliktregulation in komplementären Kinderfreundschaften analysiert (Abbildung 15.1), die bezüglich der Mutterbindung der Partner als komplementär gelten (z. B. Alisch & Wagner 2006; Altmann 2010). Die Erfassung der Konfliktregulation erfolgt über das Engagement, also das Dominanz- und Folgeverhalten im Interaktionsablauf. Relevante Bereiche zur Erfassung des Engagements sind

- die interpersonale Distanz,
- das Blickverhalten und die Kopfneigung,
- die Körperorientierung und
- die Mimik.

Zur Messung und effizienten Auswertung feinaufgelöster Engagementverlaufsdaten können auf den Daten einer 3D-Kamera basierende bewegungsanalytische photogrammetrische Verfahren zum



Einsatz kommen und die bislang etablierten Konzepte der Videografie mit einer sich anschließenden aufwendigen interaktiven Annotation ablösen (vgl. [Kapitel 12](#)). Im Folgenden werden zwei entwickelte Ansätze zur spatiotemporal hochaufgelösten Erfassung der interpersonalen Distanz ([Abschnitt 15.1.2](#)) und der Körperorientierung ([Abschnitt 15.1.3](#)) durch automatische Verfahren der photogrammetrischen Bewegungsanalyse kurz vorgestellt.



**Abb. 15.1:** Konflikte unter Kindern in Kindertagesstätten (Dörfler u. a. 2001).

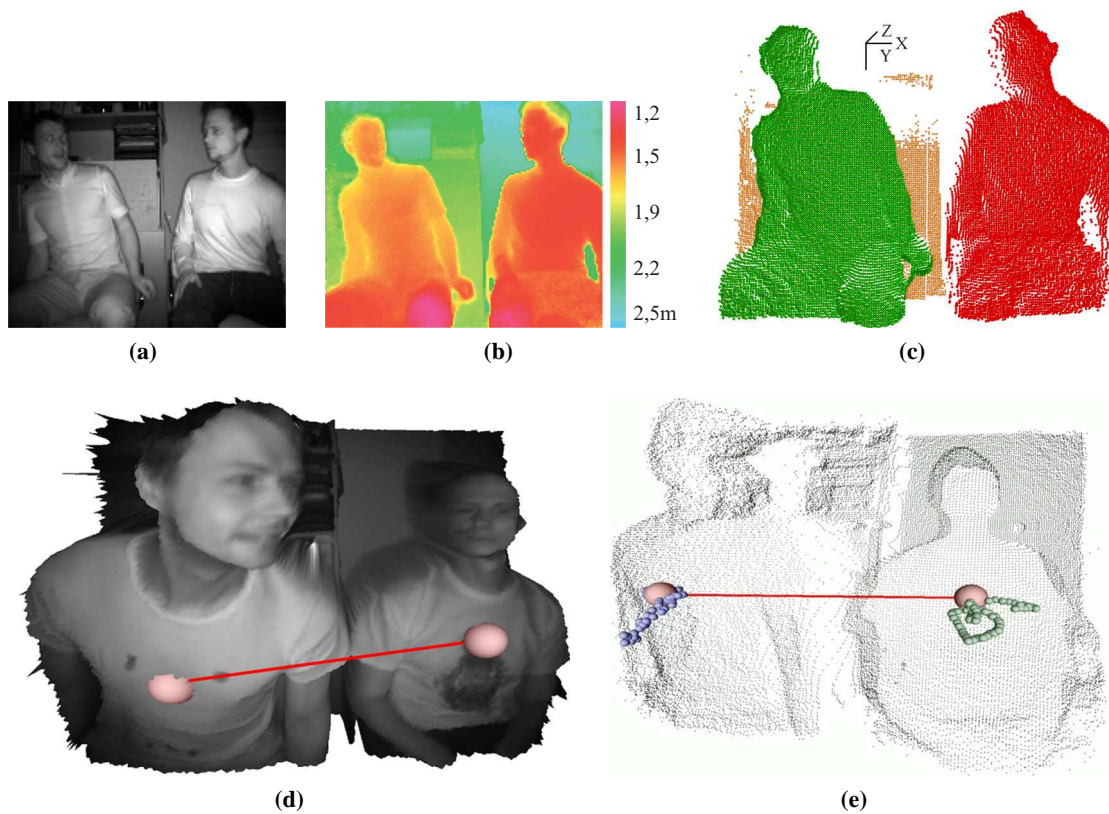
### 15.1.2 Interpersonale Distanz

Die [Abbildungen 15.2a](#) und [15.2b](#) zeigen die Eingabedaten einer 3D-Kamerabildsequenz zweier interagierender Personen. Die Hülle der Probanden wird automatisch aus dem ersten Amplituden- und Tiefenbild der Sequenz durch Schwellenwertbildung, Clustering und Konnektivitätsanalyse segmentiert ([Abbildung 15.2c](#)). Eine Erweiterung des Segmentierungsansatzes, in den neben dem Grauwert- und Tiefenbild der 3D-Kamera auch durch 2,5D-LST bestimmte dichte Geschwindigkeitsvektorfelder eingehen, ist zusätzlich denkbar. Für jede Epoche erfolgt auf der Basis des entwickelten integrierten 2,5D-LST ([Kapitel 14](#)) die Verfolgung von Körperoberflächensegmenten, welche z. B. auf die Schwerpunkte der segmentierten 3D-Punktwolken zentriert sein können ([Abbildungen 15.2d](#) und [15.2e](#)). Die aus den Körperpunkttrajektorien gerechneten interpersonalen Distanzen werden schließlich zur statistischen Auswertung im erziehungswissenschaftlichen Kontext herangezogen. [Westfeld \(2007b\)](#) gibt hierzu ausführliche Informationen.

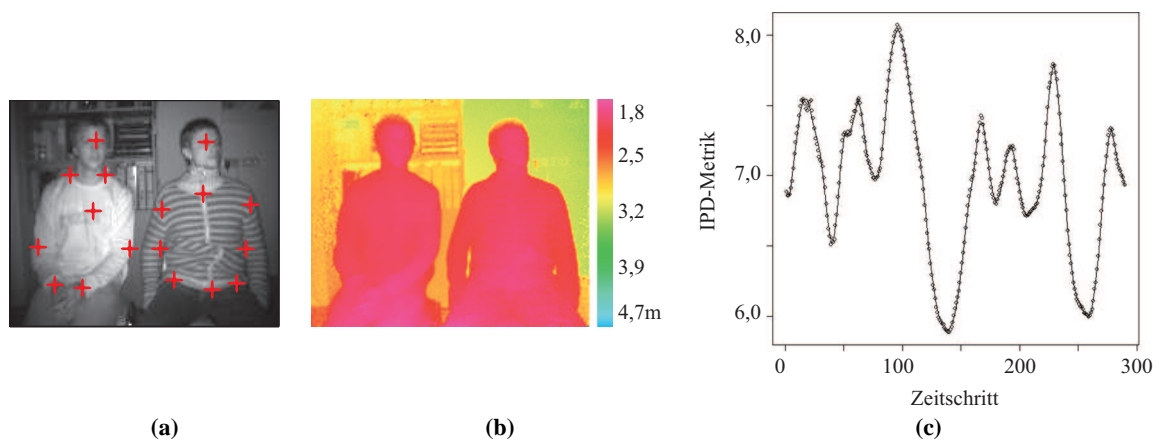
Die Reduktion der segmentierten Körperhüllen auf lediglich jeweils einen diskreten Punkt resultiert bei einem stark gestikulierenden Gesprächspartner in der Verschiebung des Körperschwerpunkts in Richtung des ausgetreckten Arms. Für den soeben beschriebenen Ansatz bedeutet dies die Notwendigkeit einer Subsegmentierung in einzelne Körperteile (wie Kopf, Rumpf, Arme und Beine) und einer Verfolgung der einzelnen Körperteilschwerpunkte.

Alternativ können aber auch für jede Person  $i$  repräsentative Körperpunkte im ersten Bild  $t = 1$  definiert und zu einem anwendungsspezifisch durch die Matrix  $\mathbf{\Omega}$  gewichteten abstrakten Körpermodell  $\mathbf{X}_{it}$  verrechnet werden ([R. Hempel & Westfeld 2009](#); [Abbildung 15.3a](#)). Für jeden weiteren Zeitpunkt  $t$  der Bildsequenz werden die Körperpunkte durch 2,5D-LST einander zeitlich zugeordnet. Die Berechnung der interpersonalen Distanz erfolgt über

$$d(\mathbf{X}_{1t}, \mathbf{X}_{2t}) := d_t := \left[ \text{spur}(\mathbf{X}_{1t} - \mathbf{X}_{2t})^T \mathbf{\Omega} (\mathbf{X}_{1t} - \mathbf{X}_{2t}) \right]^{\frac{1}{2}} \quad (15.1)$$



**Abb. 15.2:** Automatische Berechnung der interpersonalen Distanz über die Verfolgung der Körperschwerpunkte in 3D-Kameradaten (Westfeld 2007b): (a) und (b) zeigen die Eingabedaten, (c) zeigt die segmentierte 3D-Punktwolke. Die Abbildungen (d) und (e) visualisieren die Ergebnisse des integrierten 2,5D-LST als 3D-Amplitudenbild und 3D-Punktwolke mit den berechneten Trajektorien.



**Abb. 15.3:** Statistische Modellierung der interpersonalen Distanz aus RIM-Daten (R. Hempel & Westfeld 2009): Die in das Amplitudenbild (a) geplotteten repräsentativen Oberflächenpunkte bilden ein abstraktes Körpermodell. Sie werden mittels 2,5D-LST verfolgt und zur Berechnung einer IPD-Metrik herangezogen, welche in (c) dargestellt ist.

in welche abschließend die für den Erziehungswissenschaftler relevante Kurve ausgleichend eingepasst wird (Abbildung 15.3c).

### 15.1.3 Körperorientierung

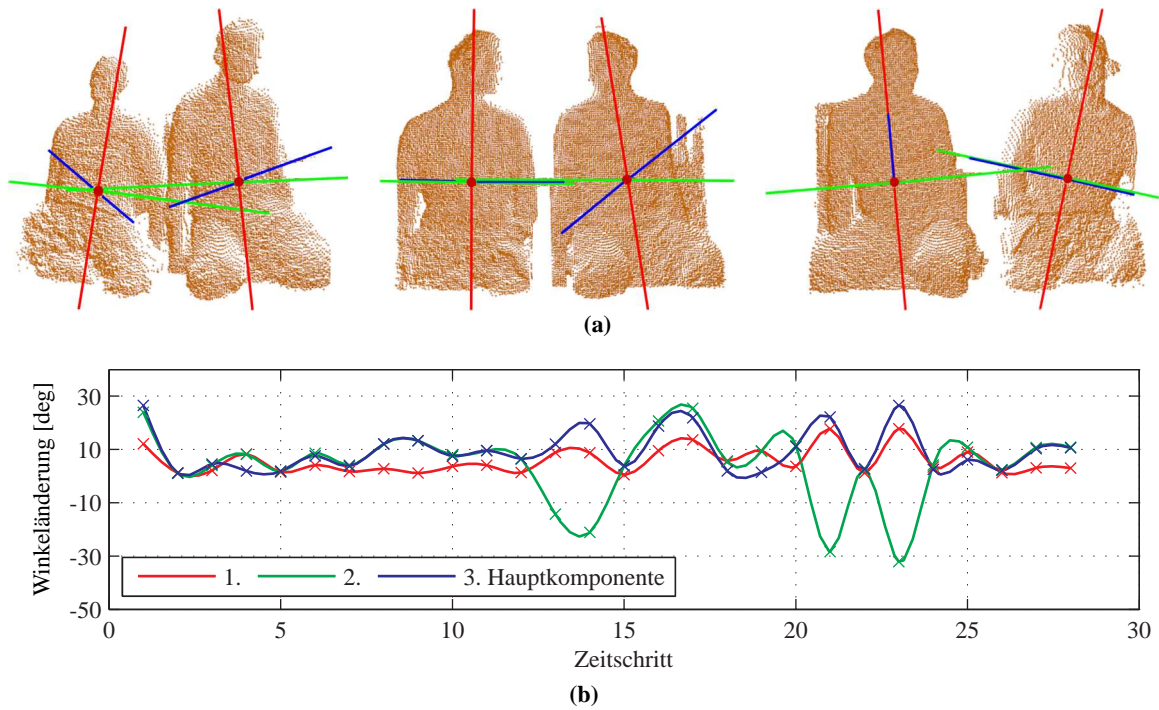
Neben der interpersonalen Distanz liefert die Körperorientierung, also die gegenseitige Zu- bzw. Abwendung beteiligter Probanden, wichtige Informationen über einen Interaktionsverlauf. Basierend auf den segmentierten Daten des Abschnitts 15.1.2 und ausgehend von dem im ersten Bild der Sequenz bestimmten und mittels 2,5D-LST zugeordneten Körperschwerpunkt wird zunächst für jede Epoche die erste Hauptkomponente in Richtung der größten Varianz gelegt. Die zweite steht rechtwinklig dazu in Richtung der zweitgrößten Varianz. Die dritte Hauptkomponente ergänzt das Modell zu einem kartesischen Koordinatensystem (Abbildung 15.4a). Diese als Hauptkomponentenanalyse (HKA) bekannte Vorgehensweise wird mathematisch in z. B. (Fukunaga 1990) beschrieben. Die nun folgende Analyse der Winkeländerungen der Geraden liefert die in Abbildung 15.4b dargestellten Kurvenverläufe, welche zur weiteren sozialpädagogischen Auswertung genutzt werden können. Da dieses Verfahren ebenfalls schwerpunktbasiert arbeitet, ist auch hier eine Subsegmentierung in der Praxis meist unerlässlich. In (Westfeld 2007b) geschieht dies bereits für den Kopf.

Eine zweite Möglichkeit besteht in der initial interaktiven oder automatischen Detektion der beiden Schulterpunkte  $S_1$  und  $S_2$  im ersten Bild und der 2,5D-LST-Verfolgung durch den weiteren Verlauf der Sequenz (Abbildung 15.5a). Unter der Annahme, dass sich ein Proband nicht um die eigene Schulterachse dreht, können so absolute und relative Winkeländerungen zeitlich hochaufgelöst erfasst und dokumentiert werden (Abbildung 15.5b). Zur Erhöhung der Zuverlässigkeit des Zuordnungsprozesses wird eine Bedingungsgleichung  $d_S = (\Delta X_S^2 + \Delta Y_S^2 + \Delta Z_S^2)^{\frac{1}{2}}$  eingeführt, welche eine zeitlich konstante Schulterbreite fordert (vgl. Abschnitt 4.4.3). Die Sollstrecke wird im ersten Bild der Sequenz initialisiert und fortlaufend nach jedem weiteren Bild durch Mittelbildung verbessert. Um gewisse Schwankungen bspw. durch Heben und Senken der Schulterblätter zuzulassen, wird die Bedingungsgleichung im Ausgleichungsprozess nicht unter vollem Zwang gewichtet.

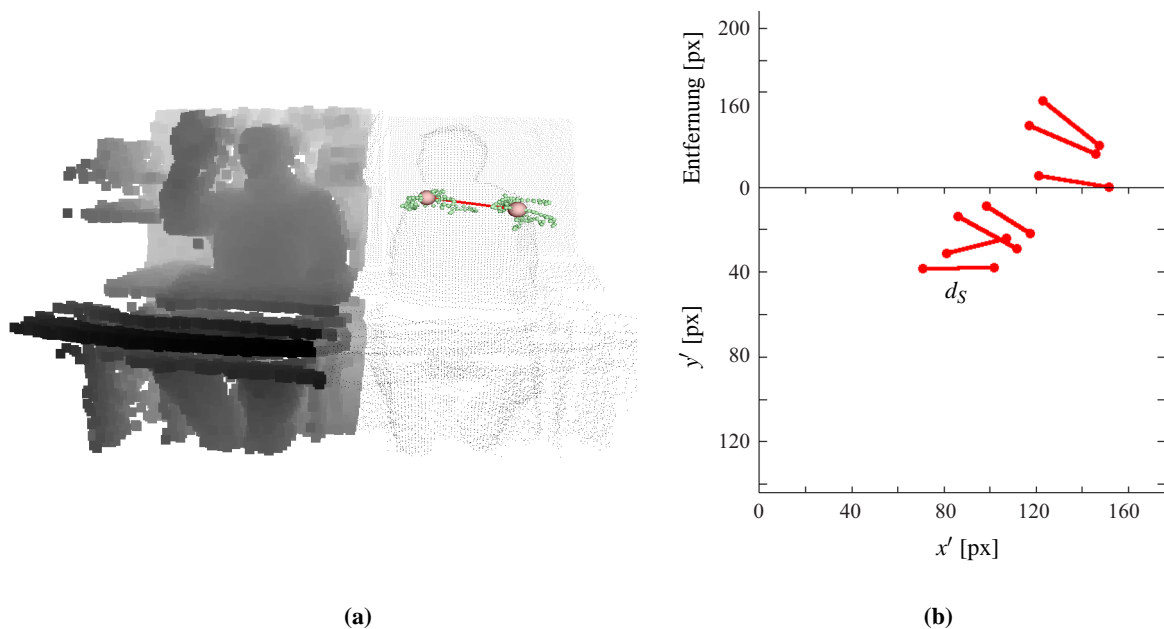
## 15.2 Gestenerkennung

### 15.2.1 Motivation

Der Schwerpunkt der Psychotherapieforschung liegt gegenwärtig in der Prozess-*Outcome*-Forschung, welche die Erhebung von Daten zum Therapieverlauf (Prozess) mit den Resultaten des Therapieergebnisses (*Outcome*) verknüpft. Das nonverbale Interaktionsverhalten zwischen Patient und Therapeut kann dabei Aufschluss über den Erfolg einer Psychotherapie geben. Speziell dient hier die videografische Analyse von Handbewegungen inklusive Gesten als neuropsychologische Grundlage von Bewegungsverhalten und dessen Störungen bei neurologischen und psychischen Erkrankungen (Abbildung 15.6a; Lausberg u. a. 2007).



**Abb. 15.4:** Automatische Berechnung der Körperorientierung aus 3D-Kameradaten (Westfeld 2007b): Ausgehend von den mittels 2,5D-LST durch die Sequenz verfolgten Körperschwerpunkten der segmentierten 3D-Kamerapunktewolken werden die Hauptkomponenten eingepasst (a) und zur Ableitung der Änderung der Körperorientierung herangezogen (b).



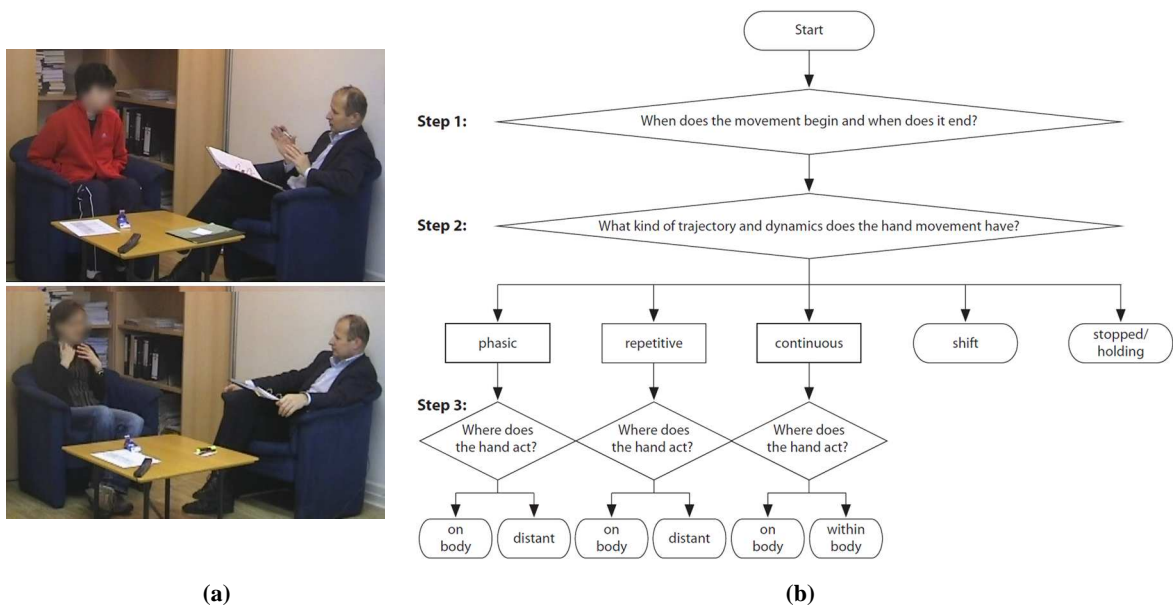
**Abb. 15.5:** Automatische Verfolgung der Schulterpunkte durch eine 3D-Kamerabilddsequenz: (a) 3D-Trajektorien (grün) der Schulterpunkte (rot). (b) Plot einzelner Schulterpunkte mit der als Restriktion formulierten Strecke  $d_s$ .



In diesem Zusammenhang entwickeln [Lausberg und Sloetjes \(2009\)](#) das Bewegungsanalyse-System **NEUROGES** (*Neuropsychological Gesture Coding System*), welches sprachbegleitende Handgesten aus neuropsychologischer Sicht klassifiziert und in die vom Max-Planck-Institut für Psycholinguistik (Nijmegen, NL) erstellte und zum freien Download angebotene Annotationssoftware *ELAN* einspeist. **NEUROGES** besteht aus drei Modulen: Modul I analysiert das jeweilige Bewegungsverhalten der linken und rechten Hand, Modul II untersucht anschließend das Verhalten der Hände zueinander und Modul III dient der funktionalen Gestenkodierung.

Die interaktive Analyse der Videos erfolgt durch insgesamt drei Spezialisten (vgl. [Abschnitt 12.1](#)). In einem ersten Arbeitsschritt markiert ein Auswerter die Bewegungseinheiten (*Tagging*). Zwei zusätzliche unabhängige *Rater* vollziehen die weiteren Kodierschritte zunächst am Patienten und dann am Arzt, ohne Kenntnis des Inhalts der Interviews sowie der Arbeitshypothesen. Besonders die Abarbeitung des ersten Moduls ([Abbildung 15.6b](#)) zur Detektion der sich bewegenden Hand, zur Bestimmung der temporalen Struktur der Bewegung und zur räumlichen Lokalisierung des Orts, an dem die Bewegung ausgeführt wird, ist sehr zeitaufwendig und kann durch automatische Verfahren der photogrammetrischen Bildsequenzanalyse auf der Basis von z. B. 3D-Kameradaten erleichtert werden.

Die folgenden [Abschnitte 15.2.2 bis 15.2.4](#) veranschaulichen das Potential einer automatisch durchgeführten 3D-Kamerabildsequenzanalyse exemplarisch anhand der Auswertung des Moduls I.



**Abb. 15.6:** (a) Videoaufnahmen von Patient-Therapeuten-Gesprächen zur Dokumentation sprachbegleitender kommunikativer Gesten. (b) Algorithmus zur kinetischen Kodierung von Gestentypen der linken und rechten Hand ([Lausberg & Sloetjes 2009](#)).

### 15.2.2 Markierung der Bewegungseinheiten

Die Identifizierung und Markierung von Bewegungseinheiten sind die initialen Schritte des Moduls I zur kinetischen Kodierung von Gestentypen. Die Bewegungseinheiten einer Hand sind definiert als die Zeiten zwischen jeweils zwei aufeinanderfolgenden Grundstellungen (*Home Base Position*; z. B. Hände im Schoß gefaltet). Sie können automatisch über die Berechnung von spatiotemporal aufgelösten 2,5D-LST-Bewegungsvektorfeldern bestimmt werden.

Die [Abbildung 15.7](#) zeigt beispielhaft die durch 2,5D-LST berechneten Trajektorien einer linken und rechten Hand über 100 3D-Kamerabilder. Deutlich zu erkennen ist die – in diesem Fall repetitive – Bewegung der linken Hand, wohingegen sich die rechte Hand in Ruhe befindet. Für die primäre Auswertung der gemessenen Handbewegungen sind zwei statisch bzw. dynamisch festgelegte Schwellenwerte  $\varepsilon_1$  und  $\varepsilon_2$  zu definieren.  $\varepsilon_1$  zeigt an, ob eine signifikante Bewegung zwischen zwei Zeitschritten vorliegt.  $\varepsilon_1$  setzt sich zusammen aus der Summe eines messspezifischen Schwellenwerts in der Größenordnung des Messrauschens und eines behavioralen Schwellenwerts, welcher aus psychologischer Sicht nicht signifikante Bewegungen berücksichtigt.  $\varepsilon_2$  legt zusätzlich die minimale Länge des Intervalls zwischen zwei Grundstellungen fest.

### 15.2.3 Klassifikation der Bewegungseinheiten

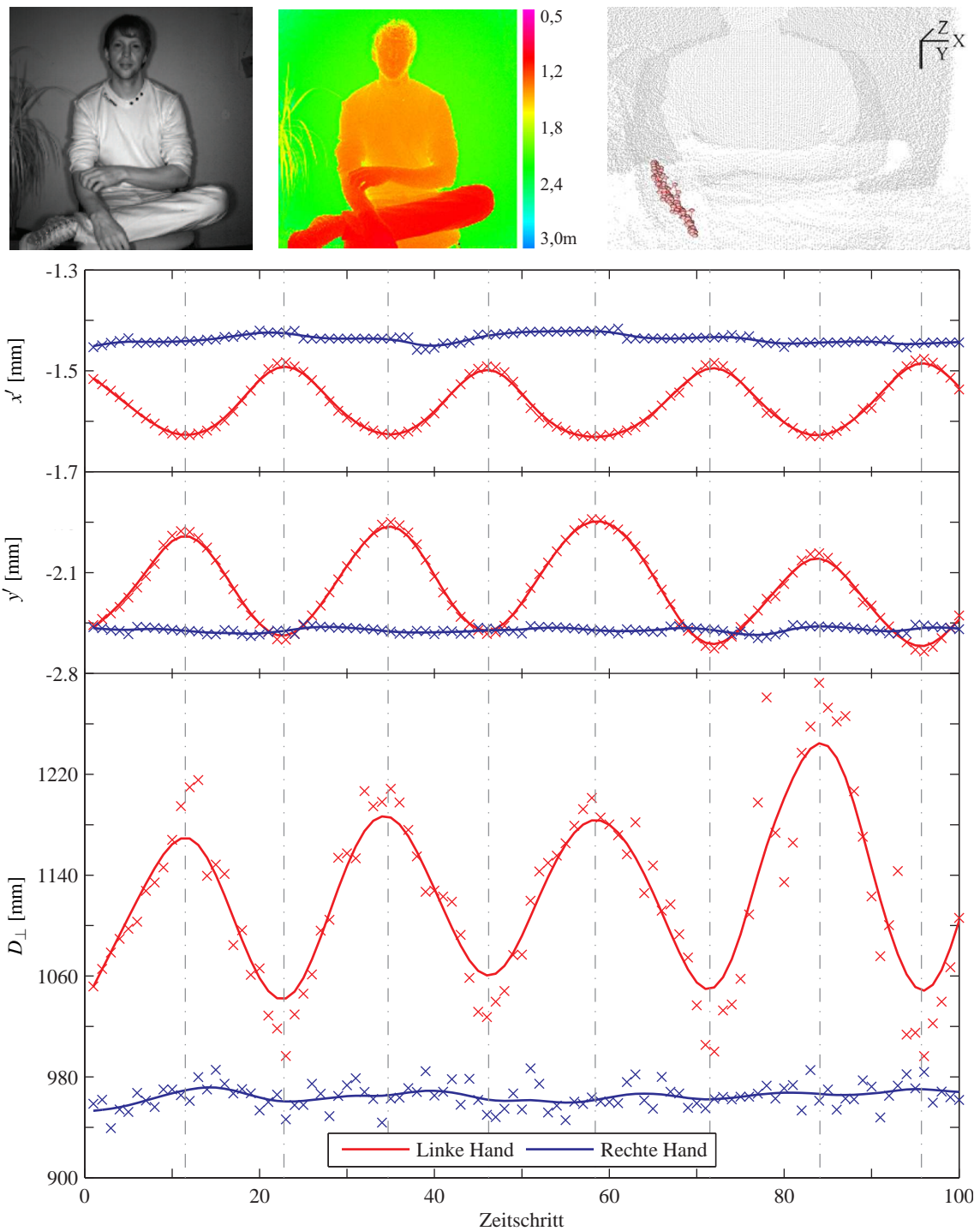
Im Anschluss sind die detektierten Handbewegungen bezüglich ihrer spatiotemporalen Bewegungsstruktur zu interpretieren. [Lausberg und Sloetjes \(2009\)](#) unterscheiden zwischen den folgenden Bewegungsarten ([Abbildung 15.6b](#)):

- Phasisch verlaufende *vor-* bzw. *vor-zurück-*Bewegung
- Repetitive *vor-zurück-vor-*Bewegung auf demselben Weg
- Lange und unregelmäßig verlaufende kontinuierliche Handbewegung
- Geradlinige Handbewegung von Grundstellung zu Grundstellung ohne weitere Gesten
- Aktives Halten einer Hand gegen die Schwerkraft

Beispielhaft für die Klassifikation der Bewegungseinheiten kann in der [Abbildung 15.7](#) eine sich wiederholende *vor-zurück-vor-*Selbstberührung der linken Hand am rechten Arm detektiert werden. Repetitive Bewegungen resultieren in periodischen  $x'$ -,  $y'$ -, und/oder Distanztrajektorien, die durch sich überlagernde Sinus-Kosinus-Funktionen höherer Ordnung analytisch modellierbar sind. Die gefundene Funktion kann anschließend hinsichtlich ihrer Perioden (Maß für die Wiederholungsanzahl) oder der Auslenkung ihrer Amplituden (Maß für die Intensität der Bewegung) untersucht werden. Ebenfalls möglich ist das Ableiten von Wartezeiten zwischen sich wiederholenden Bewegungen.

### 15.2.4 Identifizierung des Bewegungsorts

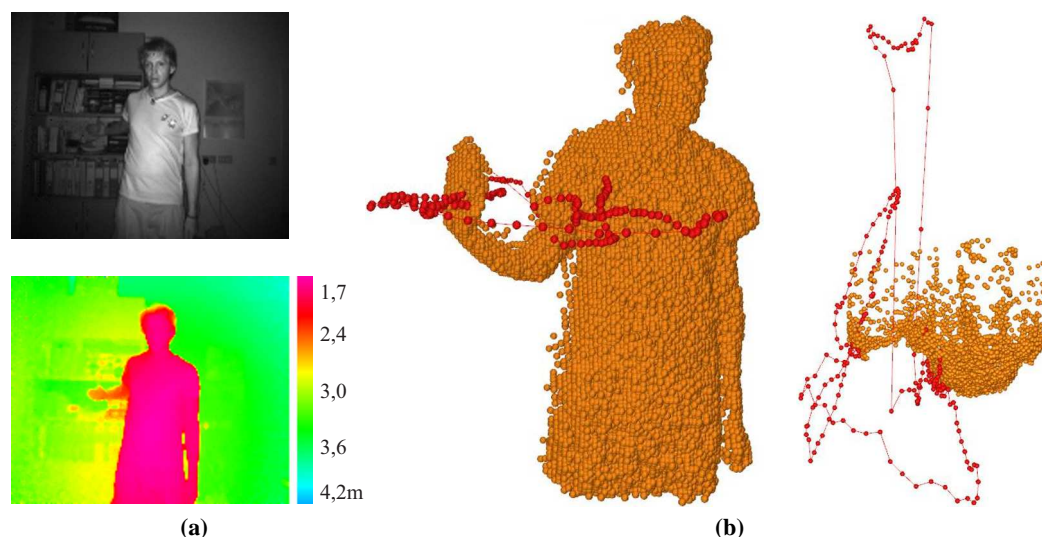
Der Bewegungsort einer sprachbegleitenden Geste ist ebenfalls von Interesse und wird laut [Lausberg und Sloetjes \(2009\)](#) in körpernah und körperfern unterschieden. Die [Abbildung 15.8a](#) stellt auch hierfür beispielhaft ein Amplituden- und farbkodiertes Entfernungsbild einer 250 Bilder umfassenden



**Abb. 15.7:** Automatische Detektion, Markierung und Klassifikation einer sprachbegleitenden Geste: Oben: 3D-Kameraeingabedaten und 3D-Punktwolkenausschnitt mit dem Trajektorienplot der linken Hand. Unten: Dargestellt werden die mittels 2,5D-LST berechneten  $x'$ -,  $y'$ - und  $D_{\perp}$ -Trajektorien beider Hände. Die linke Hand führt eine sich wiederholende Selbstberührung aus, die rechte befindet sich in der Grundstellung in Ruhe.



Sequenz dar. Die [Abbildung 15.8b](#) zeigt den 3D-Trajektorienverlauf einer mit 2,5D-LST verfolgten rechten Hand. Die Zeitintervalle körpernaher bzw. körperberührender Positionen sind deutlich zu erkennen. Werden neben den Händen zusätzliche Punkte zur Definition der Körperposition im Raum verfolgt, kann über eine einfache dynamische Schwellenwertanalyse in allen drei Koordinatenrichtungen der Bewegungsort relativ zum Probanden identifiziert werden.



**Abb. 15.8:** 3D-Identifizierung des Bewegungsorts einer Geste: (a) Eingabedaten. (b) Frontansicht und Draufsicht einer 3D-Punktwolke, überlagert mit dem Trajektorienverlauf der rechten Hand.

## 15.3 Diskussion

Das [Kapitel 15](#) zeigte am Beispiel der menschlichen Bewegungs- und Verhaltensanalyse das Potential, das in einer 3D-Kameradatenaufnahme und -auswertung liegt. Unter der Berücksichtigung system- und prinzipbedingter limitierender Faktoren, wie die geringe räumliche Auflösung des Sensors, die eingeschränkte Reichweite der Infrarotdioden oder das [Signal-Rausch-Verhältnis \(SNR\)](#), stellen 3D-Kameras bereits jetzt eine interessante Alternative zu bspw. aufwendigen multiokularen Experimentalkonfigurationen dar. Besonders vorteilhaft sind sie bei der Akquisition und Auswertung von 3D-Daten in Videorate im Bereich der noninvasiven menschlichen Bewegungsanalyse.

Die gezeigten Bildsequenzen wurden in Zusammenarbeit mit Kollegen der jeweiligen Disziplin entsprechend den in der Praxis zu erwartenden Versuchsanordnungen akquiriert. Diese Daten dienen dem entwickelten Zuordnungsverfahren 2,5D-LST zur berührungslosen und vollständig automatischen Herstellung zeitlicher Korrespondenzen anwendungsspezifisch relevanter Körperpunkte. Darauf aufbauend wurden Ansätze zur weiteren Auswertung der erhaltenen spatiotemporal hochauflösten Daten aufgezeigt. Eine umfassende anwendungsspezifische Genauigkeits- und Zuverlässigkeitsanalyse nach z. B. den Beurteilungskriterien Genauigkeit, Zuverlässigkeit, Größe des Konvergenzradius, Effektivität und Anwendungsflexibilität ([Barnard & Fischler 1982](#)) sowie eine Evaluie-

rung anhand von bspw. interaktiv gewonnenen Vergleichsdaten kann Gegenstand zukünftiger Arbeiten auf diesem Gebiet sein.



## Zusammenfassung und Ausblick



## Zusammenfassung

Die quantitative Erfassung von Bewegungsabläufen in der Analyse nonverbaler Kommunikation in Interaktionsverläufen von Personen spielt eine große Rolle auf den Gebieten der Erziehungswissenschaften und Psychologie. Die meisten der momentan eingesetzten Verfahren erheben die notwendigen Bewegungsdaten mit gebräuchlicher Videotechnik und analysieren diese anschließend interaktiv an entsprechend ausgestatteten Playbackstationen durch die Verwendung von Kodierungswerkzeugen. Das Konzept ist durch den Anwender einfach zu realisieren, birgt jedoch auch signifikante Nachteile. Vor allem im Hinblick auf eine spatiotemporal hochaufgelöste, effektive und auswertungsabhängige Erhebung und Analyse nonverbaler Verlaufsdaten besteht seit langem der Wunsch nach einer automatisierten Auswertung.

Die Aufnahme- und Auswertestrategien der Photogrammetrie und der Bildanalyse können dazu einen wertvollen Beitrag leisten. Bei einer stereoskopischen Aufnahme ist die dreidimensionale Rekonstruktion von Bewegungsabläufen aus Bildsequenzen möglich. Als Basis für eine detaillierte Bewegungsanalyse dienen Methoden der Bildverfolgung. Bisherige stereo- oder multiskopische Aufnahmesysteme werden hauptsächlich in Bereichen der Biomechanik, der Rehabilitation und der Filmanimation eingesetzt. Solche kommerziell erhältlichen Trackingsysteme beschränken sich allerdings auf die Bestimmung von 3D-Trajektorien retroreflektierender Zielmarken, welche exakt am Menschen zu positionieren sind. Das Messverfahren kann damit streng genommen nicht mehr als noninvasiv charakterisiert werden. Außerdem bringt der Einsatz mehrerer Kameras einen nicht unerheblichen sensortechnischen Aufwand und eine teils hohe Komplexität bei der Lösung raumzeitlicher Zuordnungsaufgaben mit sich.

Zur Reduktion dieses instrumentellen und verarbeitungstechnischen Aufwands der Bewegungsanalyse bieten sich Tiefenbildsensoren als interessante monosensorielle Alternative der Datenakquisition an. 3D-Kameras auf der Grundlage von Photomischdetektoren oder vergleichbaren Prinzipien stehen derzeit mit einem Sensorformat von bis zu  $204^2$  px und einer zeitlichen Auflösung von bis zu 50 Bildern pro Sekunde zur Verfügung. Sie liefern durch die simultane Aufnahme von Grauwert- und Entfernungsbildern für jedes Pixel zusätzlich zur Bildinformation einen Distanzmesswert. Die auf einem regelmäßigen Raster im Bildraum angeordneten 3D-Punkte eines mit hoher Wiederholrate aufgenommenen Oberflächenmodells können unter Kenntnis der Kamerageometrie direkt berechnet werden. Somit entfällt eine komplexe und rechenaufwendige stereoskopische Zuordnung, was bei bewegungsanalytischen Aufgabenstellungen als entscheidender Vorteil zu werten ist. Trotz der im Vergleich zu Stereokamerasystemen geringen räumlichen Auflösung besitzen 3D-Kameras großes Potential für Anwendungen der menschlichen 3D-Bewegungsanalyse, als Einzelsensoren multiokulare Systeme zu ersetzen.

Der Einsatz einer 3D-Kamera als Messinstrument erfordert die exakte Beschreibung des optischen Abbildungsvorgangs sowie der simultan durchgeführten Streckenmessung. Ein Ziel der vorliegenden Dissertation war die Sensormodellierung und die damit einhergehende Entwicklung eines Verfahrens zur integrierten Kalibrierung. Um das Potential des Tiefenbildsensors optimal ausnutzen zu können, wurden zunächst die auf eine 3D-Kameramessung wirkenden Fehlereinflüsse beschrieben und parametrisiert. Der Abbildungsvorgang einer 3D-Kamera folgt dem Modell der Zentralprojektion. Zur Modellierung optospezifischer Fehler konnte somit der den Zusammenhang zwischen Bild- und Kamerakoordinatensystem beschreibende und in der Photogrammetrie übliche Parametersatz der inneren Orientierung übernommen werden. Ein Korrekturmodell zur streng geometrisch-physikalischen Modellierung distanzmessspezifischer Fehlereinflüsse wurde aus der Literatur und eigenen Genauigkeitsuntersuchungen abgeleitet. Für (noch) nicht parametrisierte Fehlerquellen sind Vorschläge zur Elimination oder teilweisen Reduktion gegeben worden.

Die auf dieser Fehlerbetrachtung aufbauende und im Rahmen der Arbeit entwickelte integrierte Bündelblockausgleichung verknüpft in ihrem funktionalen Modell zwei gemeinsam auszugleichende direkt in beiden Informationskanälen beobachtete Größen: die 2D-Mittelpunktkoordinaten von in Amplitudenbildern kreisförmig abgebildeten Referenzkugeln und die Distanzen zu aus den Entfernungsbildern segmentierten 3D-Punkten der korrespondierenden Kugeloberflächen. Zusätzlich eingeführte geometrische Nebenbedingungen stabilisieren das Gleichungssystem. Die stochastischen Zusammenhänge des heterogenen Beobachtungsmaterials wurden statistisch adäquat über die simultane Schätzung der Varianzkomponenten berücksichtigt (**Varianzkomponentenschätzung, VKS**). Als Ergebnis liegen neben den äußeren Orientierungsparametern und den Objektkoordinaten der Punkte der Kugelmitten und -oberflächen die für den weiteren Verlauf relevanten Werte zur Korrektur der Bilddaten und Distanzmessungen vor. Als statistisches Verfahren liefert die integrierte Bündelblockausgleichung darüber hinaus Informationen zur Quantifizierung des Genauigkeitspotentials der ursprünglichen und ausgeglichenen Beobachtungen sowie der geschätzten Unbekannten. Der gesamte Programmablauf des auf einer Selbstkalibrierung basierenden Ansatzes bedarf keines Nutzereingriffs; sowohl die Bereitstellung der Beobachtungen als auch die Beschaffung von Näherungswerten für die unbekannt Parameter vollzieht sich voll automatisch.

Die integrierte Bündelblockausgleichung wurde anhand von Aufnahmen eines entworfenen Referenzfelds ausführlich validiert. Die durchgeführten Analysen zur Beurteilung der inneren Genauigkeit dokumentieren anhand der Residuen sowie der a-posteriori Standardabweichungen der Beobachtungen zum einen die Richtigkeit der aufgestellten Zusammenhänge, zeigen zum anderen aber auch die Schwächen des zum gegenwärtigen Zeitpunkt noch nicht alle Einflüsse berücksichtigenden Distanzkorrekturmodells. Dessen flexible Erweiterung ist durch die adaptive Implementierung der integrierten Bündelblockausgleichung zukünftig jedoch problemlos möglich. Ein Vergleich mit Referenzdaten übergeordneter Genauigkeit erlaubt darüber hinaus die Angabe der resultierenden Objektpunktgenauigkeit mit 5 mm in allen drei Koordinatenrichtungen. Des Weiteren zeigte die vergleichende Gegenüberstellung, dass das geschlossene Modell des vorgestellten simultanen Verfahrens durch die implizite Berücksichtigung der Wechselwirkungen aller Modellparameter die geometrisch-

physikalische Wirklichkeit besser widerspiegelt, als ein lediglich quasi-simultan arbeitender Ansatz.

Weitere Fehlerarten beeinflussen die Distanzmessung. Sinnvoll ist die Integration adaptiver Varianzkomponenten zur Erfassung linear von der Distanzmessung selbst abhängender Fehleranteile. Eine Experimentalkonfiguration, die um 3D-Kameraaufnahmen unterschiedlicher Modulationsfrequenzen und Integrationszeiten erweitert wurde, kann ebenso Gegenstand weiterführender Arbeiten sein wie eine Untersuchung zur zeitlichen Stabilität der Kalibrierparameter oder ein Vergleich der Kalibrierergebnisse verschiedener Tiefenbildsensoren.

Die Nutzung von 3D-Kameras in der 3D-Bewegungsanalyse bedingt die Adaption, Weiter- und Neuentwicklung von Verfahren der Bildanalyse. Im Rahmen der vorliegenden Arbeit wurde im Hinblick auf die Bestimmung raumzeitlich aufgelöster Interaktionsverläufe zum Zweck der Analyse menschlichen Verhaltens eine integrierte Zuordnungsmethode entwickelt, die auf die Bildsequenzen einer 3D-Kamera angewendet werden kann. *2,5D Least Squares Tracking (LST)* basiert auf der in der Photogrammetrie bekannten Bildzuordnung nach der Methode der kleinsten Quadrate (*Least Squares Matching, LSM*), welche in den 80er Jahren ursprünglich als ein Verfahren zur subpixelgenauen Bestimmung zweidimensionaler Informationen aus (Luft-)Bilddaten entwickelt wurde. Es bildet kleine Oberflächensegmente konsekutiver 3D-Kameradatenätze aufeinander ab.

Die Abbildungsvorschrift wurde auf einer 2D-Affintransformation aufbauend an die Datenstruktur einer 3D-Kamera angepasst. Für eine vollständige Ausnutzung des gesamten 3D-Kamerainformationshaushalts wurden zur Aufstellung des funktionalen Modells sowohl Grau- als auch Entfernungswerte integriert und die Parametrisierung zur Verknüpfung beider Arten von Beobachtungsgleichungen geschlossen formuliert. Neben den affinen Parametern zur Erfassung von Translations- und Rotationseffekten modellieren die Maßstabs- sowie Neigungsparameter perspektivbedingte Größenänderungen des Bildausschnitts, verursacht durch Distanzänderungen in Aufnahme- richtung. Dabei mussten folglich auch die Korrelationen zu den neu eingeführten linearen Distanzkorrekturtermen analytisch beachtet werden. Eine implementierte VKS berücksichtigte die stochastischen Eigenschaften des heterogenen Beobachtungsmaterials und schätzte die Varianzkomponenten zu jeder Iteration, was sich vor allem aufgrund der, maßgeblich von der Experimentalkonfiguration abhängenden, unterschiedlichen Genauigkeitsniveaus als sinnvoll erwies. Die Eingabedaten sind in einem Vorverarbeitungsschritt mit Hilfe der entwickelten Kalibrierroutine um ihre opto- und distanzmessspezifischen Fehler korrigiert sowie die gemessenen Schrägstrecken auf Horizontaldistanzen reduziert worden. Eine reine Amplitudenbilddauswertung liefert laterale Verschiebungsvektorfelder mit interpolierten Tiefenwerten. Im Gegensatz dazu berechnet *2,5D-LST* als integrierter Ansatz vollständige 3D-Verschiebungsvektoren. Die aus der Fehlerrechnung resultierenden Genauigkeits- und Zuverlässigkeitsangaben können als Entscheidungskriterien für die Integration in einer anwendungsspezifischen Verarbeitungskette Verwendung finden.

Bei der Validierung des Zuordnungsverfahrens wurde insbesondere untersucht, ob sich durch die Verwendung komplementärer Entfernungsinformationen Verbesserungen in der perspektivbedingten Anpassung der Größe und der Neigung des Suchbildausschnitts ergeben und sich diese ggf. positiv auf den gesamten Zuordnungsprozess auswirken. Auch der Einfluss der beiden neu eingeführten



Neigungsparameter wurde bewertet. Zusammengefasst ist zu konstatieren, dass die Einführung komplementärer Informationen eine genauere und zuverlässigere Lösung des Korrespondenzproblems bringt, vor allem bei schwierigen Kontrastverhältnissen in einem Kanal. Die Genauigkeit der direkt mit den Distanzkorrekturtermen verknüpften Maßstabs- und Neigungsparameter verbesserte sich deutlich. Darüber hinaus brachte die Erweiterung des geometrischen Modells insbesondere bei der Zuordnung natürlicher Oberflächensegmente signifikante Vorteile, die trotz ihrer geringen Ausmaße die von 2D-LSM geforderte Ebenheitsbedingung nicht erfüllen.

Es zeigte sich, dass das verwendete lineare radiometrische Modell zur a-priori Angleichung der Grauwertbildausschnitte auf die Charakteristik einer aktiven Beleuchtung mit **nahem Infrarotlicht (NIR)** angepasst werden muss. Weiterhin bleibt die Konzeptionierung eines Verfahrens zur Bestimmung der äußeren Genauigkeit sowie der interessante Vergleich mit dem auf dem Optischen Fluss basierenden und aus der *Computer Vision* stammenden 3D-Kamerazuordnungsverfahren *Optical Range Flow* zukünftigen Arbeiten vorbehalten.

Das Potential des 3D-Kamerazuordnungsansatzes wurde an zwei Anwendungsszenarien der non-verbalen Interaktionsforschung demonstriert. 2,5D-LST kam zur Bestimmung der interpersonalen Distanz und Körperorientierung im erziehungswissenschaftlichen Untersuchungsgebiet der Konfliktregulation befreundeter Kindespaare ebenso zum Einsatz wie zur Markierung und anschließenden Klassifizierung von Bewegungseinheiten sprachbegleitender Handgesten. Die Implementierung von 2,5D-LST in die vorgeschlagenen Verfahren ermöglichte eine automatische, effektive, objektive sowie zeitlich und räumlich hochaufgelöste Erhebung und Auswertung verhaltensrelevanter Daten.

Die vorliegende Dissertation schlägt die Verwendung einer neuartigen 3D-Tiefenbildkamera zur Erhebung menschlicher Verhaltensdaten vor. Sie präsentiert sowohl ein zur Datenaufbereitung entwickeltes Kalibrierwerkzeug als auch eine Methode zur berührungslosen Bestimmung dichter 3D-Bewegungsvektorfelder. Die Arbeit zeigt, dass die Methoden der Photogrammetrie auch für bewegungsanalytische Aufgabenstellungen auf dem bisher noch wenig erschlossenen Gebiet der Verhaltensforschung wertvolle Ergebnisse liefern können. Damit leistet sie einen Beitrag für die derzeitigen Bestrebungen in der automatisierten videographischen Erhebung von Körperbewegungen in dyadischen Interaktionen.

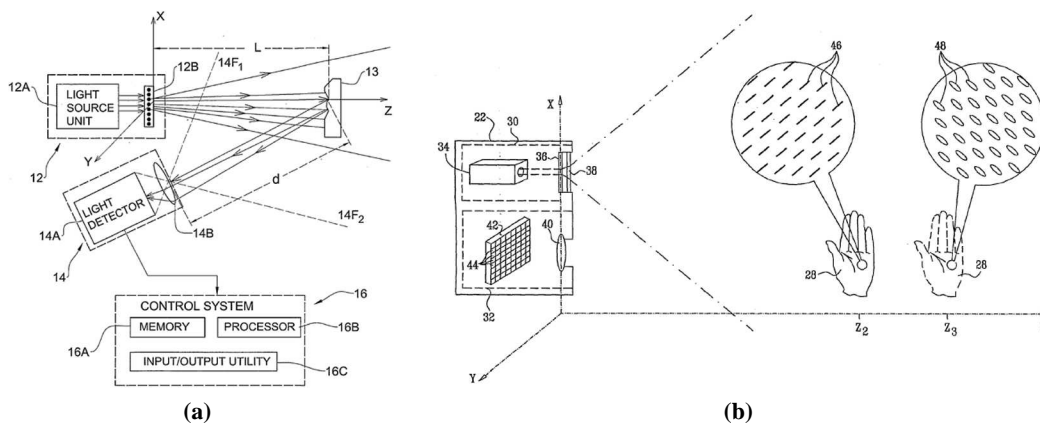
## Ausblick

Die Entwicklung auf dem Gebiet der Tiefenbildgenerierung ist rasant. Neben der in der vorliegenden Arbeit genutzten ToF-Tiefenbildmesstechnik (*Time-of-Flight*) sorgt die Vorstellung einer Hardware zur Bewegungssteuerung von Konsolenspielen für einen enormen Innovationsschub. Die US-amerikanische Microsoft Cooperation sicherte sich durch den Einkauf der israelischen Start-up-Unternehmen 3DV Systems Ltd. und PrimeSense Ltd. die Kompetenzen im Bereich der berührungslosen Mensch-Computer-Interaktion (*Human-Computer Interaction, HCI*; Abschnitt 8.4). Die beteiligten Firmen entwickelten gemeinsam den Kinect-Sensor, ein System, welches den Spielern der Videospielekonsole Xbox 360 die Steuerung der Software allein durch Körperbewegungen ermöglicht (Abbildung 17.1). Seit der weltweiten Markteinführung im November 2010 verkaufte sich die Sensorleiste millionenfach (sueddeutsche.de 2011). In Deutschland kostet sie derzeit weniger als € 100. Die Firma Microsoft verfolgt primär die umsatzsteigernde Aufwertung ihrer Videospielekonsole Xbox 360 und gab in Konsequenz die Schnittstellen zur Anwendungsprogrammierung (API) zunächst nicht frei. Der enorme Zuspruch für diese vor allem aufgrund der geringen Anschaffungskosten massenmarktauglichen Technologie zur 3D-Datengewinnung sowie das Erscheinen von inoffiziellen Treibern für Windows 7 und Linux ließen Microsoft jedoch rasch einlenken. Im Sommer 2011 wurden mit dem *Kinect for Windows Software Development Kit (SDK)* die Datenschnittstellen dokumentiert und eine Sammlung an Programmierwerkzeugen zur anwendungsbasierten Softwareentwicklung veröffentlicht. Dies soll akademischen Einrichtungen und „Hobby-Programmierern“ die Entwicklung von Applikationen für nichtkommerzielle Zwecke ermöglichen, welche über die ursprünglich fokussierte Bewegungssteuerung hinausgehen. Nicht zuletzt im Hinblick auf eine mögliche Integration in ein neues Microsoft Windows Betriebssystem wurde dies PR-wirksam als „inspirierende Nutzung“ der offenen Schnittstelle bezeichnet (Shannon Loftis, Microsoft Abteilung für Spielentwicklung).

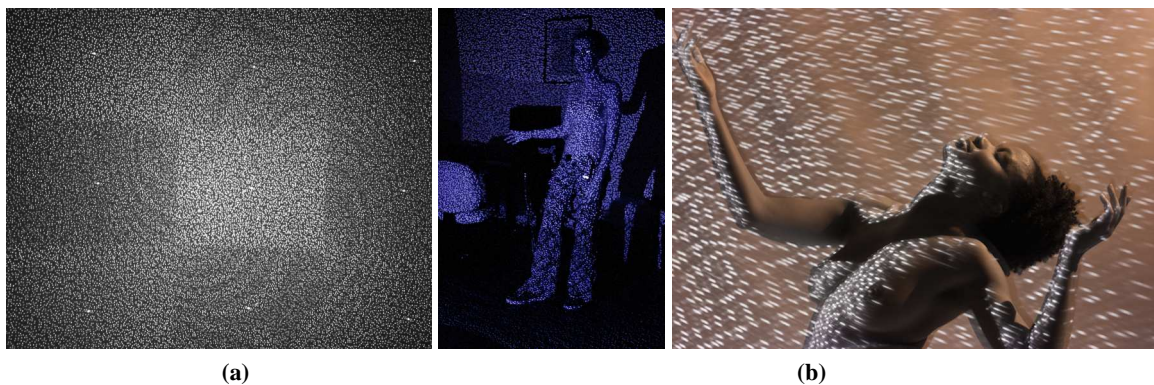
Die Kinect-Sensorleiste besteht aus einem Infrarotprojektor, einer Infrarotkamera, einer VGA-Kamera und vier Richtmikrofonen. Die Tiefeninformationen werden nicht über Modulationstechniken, sondern nach dem Prinzip der aktiven Triangulation gewonnen (Abbildung 17.2a; Abschnitt 6.2.1). Eine 60mW starke NIR-Lichtquelle (nahes Infrarotlicht; 830nm) beleuchtet den Objektraum strukturiert mit einem für jeden Zeitpunkt eindeutigen pseudo-zufälligen Specklemuster (Abbildung 17.3). Für einen Aufnahmebereich von 0,8m bis 3,5m wird das projizierte statische Muster auf einen (1200 × 900) px großen monochromen CMOS-Bildsensor (*Complementary Metal Oxide Semiconductor*) zurückreflektiert und in einer Frequenz von 30Hz registriert. Ein korrelationsbasierter Vergleich des verzerrt detektierten Specklemusters mit zuvor unter bekannten Aufnahmeabständen erfassten Referenzbildern ermöglicht die Berechnung eines 11 bit-Tiefenbilds pro Epoche (Abbildung 17.2b; Abschnitt 7.2). Die relative Orientierung zwischen Sender und Empfänger wird



**Abb. 17.1:** Kinect-Sensorleiste (Soules & Wiens 2011): NIR-Projektor, RGB-Kamera, NIR-Kamera (v. l. n. r.).



**Abb. 17.2:** Funktionsprinzip des Kinect-Sensors: (a) Die Lichtquelle sendet ein NIR-Specklemuster aus, welches von der Sensoreinheit ortsabhängig registriert wird (Zalevsky u. a. 2007). (b) Der Vergleich mit einem Referenzmuster erlaubt im Anschluss eine Rekonstruktion der erfassten Objektfläche (Freedman u. a. 2010).



**Abb. 17.3:** (a) Der Infrarotprojektor emittiert kontinuierlich ein pseudo-zufälliges NIR-Specklemuster, bestehend aus neun Grundsegmenten unterschiedlicher Helligkeit (Konolige & Mihelich 2010). (b) Der Künstler Penven (2011) verwendet das NIR-Punktmuster als Lichtquelle für seine Fotos.

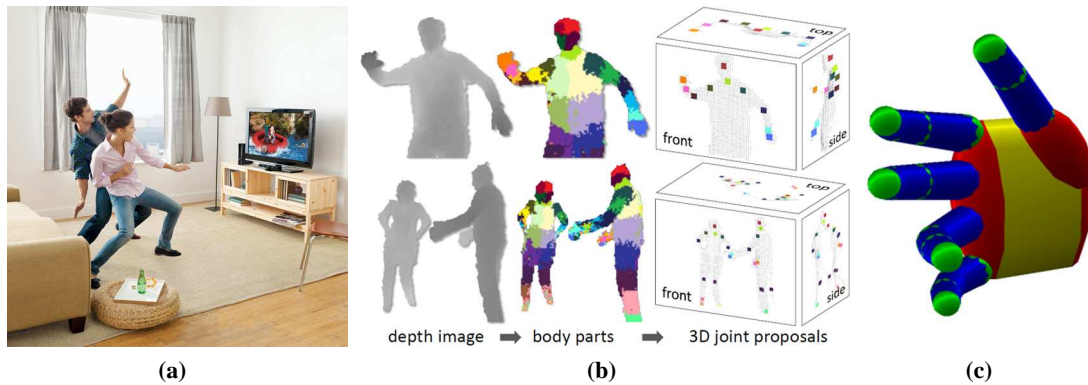
als bekannt vorausgesetzt. Die Triangulationsbasis beträgt 7,5 cm. Der Kinect-Sensor ist zusätzlich mit einer VGA-Farbkamera ausgestattet und kann somit simultan Farbbilder in einer Auflösung von  $(640 \times 480)$  px@60Hz bereitstellen. Die Wissensdatenbank [OpenKinect \(2011\)](#) bietet weitere ausführliche Informationen.

Die aus geodätisch-photogrammetrischer Sicht notwendigen Untersuchungen hinsichtlich der Eignung des Kinect-Sensors als Messinstrument sind bis jetzt überschaubar: In den Aufsätzen von [Khoshelham \(2011\)](#) und [Wujanz u. a. \(2011\)](#) werden erste Ergebnisse zu Genauigkeits- und Zuverlässigkeitsuntersuchungen sowie Verfahren zur Kalibrierung und Registrierung vorgestellt. Der Hauptfokus der veröffentlichten Beiträge liegt in der Präsentation entwickelter Methoden zur Körpererfassung, -verfolgung und -rekonstruktion zum Zweck der Steuerung von Konsolenspielen ([Schwarz u. a. 2010](#); [Oikonomidis u. a. 2011](#); [Shotton & Sharp 2011](#); [Abbildung 17.4](#)). Bereits nach kurzer Zeit wurden unzählige weitere Anwendungsszenarien präsentiert, wobei deren Spektrum aufgrund der geringen Sensorkosten sowie der frei zugänglichen Programmierschnittstellen im Vergleich zu den gezeigten ToF-3D-Kameraanwendungen bedeutend größer ist. Sie reichen von der automatischen 3D-Umgebungsmodellierung ([Henry u. a. 2010](#); [Abbildung 17.5](#)) oder der autonomen Steuerung von Robotern ([Thierfelder u. a. 2011](#); [Abbildung 17.6a](#)) über die Erfassung von Freiformkörpern ([Wujanz u. a. 2011](#); [Abbildung 17.6b](#)) bis hin zur Telemedizin ([Gallo u. a. 2011](#); [Abbildung 17.6c](#)).

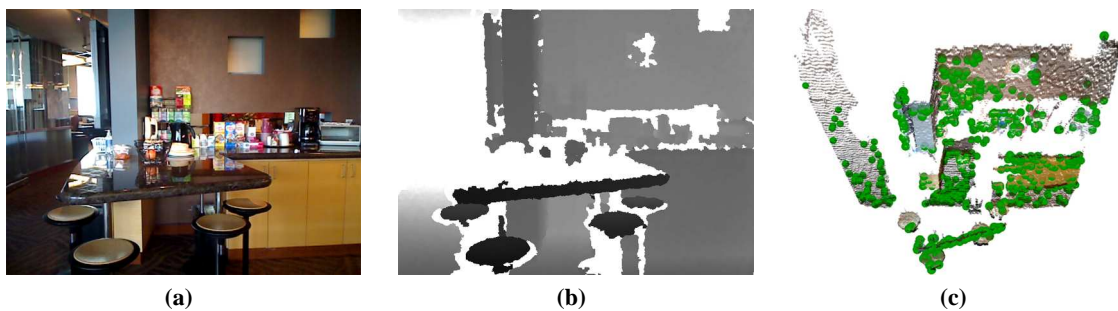
Abschließend sollen die im Rahmen der vorliegenden Dissertationsschrift entwickelten Methoden und Verfahren zur Sensorkalibrierung und Bildsequenzanalyse im Kontext der Kinect-Technologie eingeordnet werden. Die Verwendung eines Sensors als Messinstrument bedingt stets die Analyse und Modellierung von die Messung beeinflussenden Fehlern. Der funktionale Zusammenhang der entwickelten integrierten Bündelblockausgleichung beruht auf direkten Streckenmessungen, was einer direkten Übertragung auf das aktive Triangulationsprinzip widerspricht. Die Grundidee einer integrierten Kalibrierung unter der Verwendung aller zur Verfügung stehenden Beobachtungen zur vollständigen Ausnutzung des Informationshaushalts kann für die Kinect als RGBD-Kamera aber übernommen werden. Eine Adaption der vorgestellten Zuordnungsmethode 2,5D-LST ist unter der Voraussetzung exakt registrierter RGB- und Tiefenbilddaten problemlos möglich. 2,5D-LST kann so einen wertvollen Beitrag zur genauen und zuverlässigen Tiefenbildsequenzanalyse leisten, wobei dieser nicht nur auf die Auswertung zwischenmenschlichen Verhaltens beschränkt sein muss. Denkbar sind auch der Einsatz zur automatischen Umgebungskartierung oder die Lösung statischer Zuordnungsaufgaben bei der Suche homologer Oberflächensegmente zu Zwecken der Punktwolkenregistrierung im Bereich der Objekterfassung und -dokumentation.

Die Verbesserungsmöglichkeiten für die entwickelten Methoden wurden bereits im [Kapitel 16](#) dokumentiert. Für die verwendeten und relativ preisintensiven 3D-Kameras bleibt zu bemerken, dass deren Messdaten noch immer einem schwer prädizierbaren Rauschen unterliegen. Darüber hinaus sind Schwächen in der Dokumentation der Hard- und Software auszuräumen sowie prinzipbedingte Probleme wie *Scattering*, Hintergrundstrahlung, Temperaturschwankungen oder Bewegungsartefakte zu lösen. Auch ist eine verbesserte interne Kalibrierung wünschenswert, um zumindest die leicht zu korrigierenden systematischen Fehleranteile in Größenordnungen mehrerer Dezimeter hersteller-

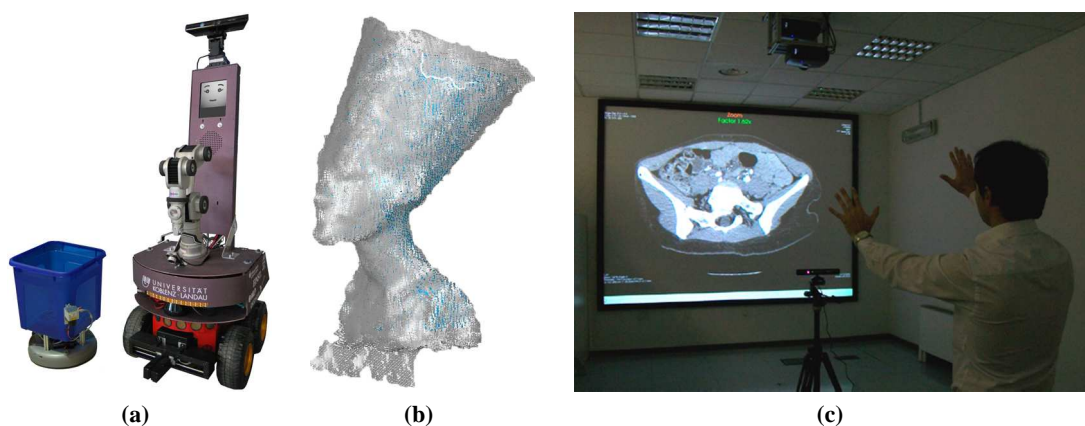




**Abb. 17.4:** Die Kinect ermöglicht die berührungslose Steuerung von Konsolenspielen: (a) „Du bist der *Controller*“ – Das Videospiel reagiert auf die erfassten Bewegungen der Spieler (Microsoft 2011). (b) 3D-Rekonstruktion einzelner Körperregionen (Shotton & Sharp 2011). (c) Detaillierte Modellierung einer Hand (Oikonomidis u. a. 2011).



**Abb. 17.5:** Daten des Kinect Sensors (Henry u. a. 2010): (a) RGB-Bild. (b) Tiefenbild. (c) Texturierte 3D-Punktwolke.



**Abb. 17.6:** (a) Die Kinect wurde beim RoboCup @ Home World Championship zur Umgebungskartierung auf dem Schwenkarm eines Roboters montiert (Thierfelder u. a. 2011). (b) 3D-Modell der Nofretete (Wujanz u. a. 2011). (c) Anwendung in der Telemedizin (Gallo u. a. 2011).

seitig zu modellieren. „Konkurrenz belebt das Geschäft“ – die oben diskutierte große Aktualität des Themas verspricht hier rasche sensortechnische Verbesserungen.

Die Entwicklungen im Bereich der Akquisition dreidimensionaler Oberflächendaten in Videorate sind sehr zukunftssträchtig und bleiben auch aus wissenschaftlicher Sicht für den Geodäten und Photogrammeter weiter spannend. Durch die Veröffentlichung der *Kinect for Windows SDK* sowie dem Aufbau einer *OpenKinect Processing Library* werden fast täglich neue Ideen und Anwendungen online auf z. B. ([KinectHacks 2011](#)) gesammelt. Nach dem Vorbild der einst von der Intel Corporation veröffentlichten *Image Processing Library* (Vorgänger der *Open Source Computer Vision Library*, OpenCV) als Innovationsbeschleunigung im Bereich der Bilderkennung kommt es so zu einem gewaltigen Entwicklungsschub in der Tiefenbildtechnologie, von dem auch die Photogrammetrie profitieren kann. Eine Integration der Kinect als sehr preiswerter 3D-Sensor bietet sich zur Entwicklung effizienter Messanordnungen und Auswertestrategien für neuartige Messaufgaben förmlich an.





## Literaturverzeichnis

- Abdel-Aziz, Y. I. & Karara, H. M. (1971). Direct Linear Transformation from Comparator Coordinates Into Object Space Coordinates in Close-range Photogrammetry. In *Proceedings of the symposium on closerange photogrammetry* (Bd. 1, S. 1-18). VA: American Society of Photogrammetry. 58
- Abramowitz, M. & Stegun, I. A. (1964). *Handbook of Mathematical Functions*. Dover Publ Inc. (Tenth Printing, December 1972, with corrections) 158, 210
- Ackermann, F. (1984). High precision digital image correlation. In *Proceedings of the 39th Photogrammetric Week* (Bd. 9, S. 231-243). 94, 97
- Ackermann, F. & Schade, H. (1993). Application of GPS for aerial triangulation. *Photogrammetric Engineering and Remote Sensing*, 59 (11), 1625-1632. 63
- Alisch, L.-M. (2010). Multimodale Situationsanalyse (MSA). Entwicklungsstand und Probleme. *Lehrerbildung auf dem Prüfstand [Sonderheft: Lehrerprofessionalität - Was wir wissen und was wir wissen müssen]*, 3, 71-85. 19
- Alisch, L.-M. & Wagner, J. W. L. (2006). *Freundschaften unter Kindern und Jugendlichen: Interdisziplinäre Perspektiven und Befunde* (1. Aufl.; L.-M. Alisch & J. W. L. Wagner, Hrsg.). Weinheim, D: Juventa. 225
- Altmann, U. (2010). Beziehungsregulation in Kinderfreundschaften – eine Prozessstudie zu Geschlechterunterschieden. In M. Haring, O. Böhm-Kasper, C. Rohlfes & C. Palentien (Hrsg.), *Freundschaften, Cliques und Jugendkulturen* (S. 105-122). VS Verlag für Sozialwissenschaften. 19, 20, 225
- Altmann, U. (2012). *Synchronisation nonverbaler Verhaltens - Weiterentwicklung und Anwendung zeitreihenanalytischer Identifikationsverfahren*. Dissertation, Fakultät für Sozial- und Verhaltenswissenschaften, Friedrich-Schiller-Universität Jena, Jena, D. 19, 20
- Altmann, U., Hermkes, R. & Alisch, L.-M. (2007). Analysis of nonverbal involvement in dyadic interactions. In A. Esposito, E. Keller & M. Faundez-Zanuy (Hrsg.), *Verbal and Nonverbal Communication Behaviours, Lecture Notes in Computer Science* (S. 37-50). Springer-Verlag, Berlin. 19, 178, 179
- Altmann, U., Hermkes, R. & Westfeld, P. (2008). *Halbautomatisierte Interaktionsforschung via Video Frame Analyse: Design, Erhebung, Analyse* [Workshop, Kulturen der Bildung, 21. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE)]. Dresden, D. 21
- Aschersleben, G. & Henning, A. (2008). Mutter-Kind-Interaktion und sozial-kognitive Entwicklung. *Magazin Forschung (Universität des Saarlandes)*, 1, 20-23. 19
- Baarda, W. (1968). *A Testing Procedure for use in Geodetic Networks* (Bd. 2) (Nr. 5). Delft:

- Netherlands Geodetic Commission. 42, 43
- Ballard, D. & Brown, C. (1982). *Computer Vision*. Prentice Hall. 91
- Baltsavias, E. P. (1988). Hierarchical multiphoto matching and DTM generation. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. 27, S. 476ff). 96
- Baltsavias, E. P. (1991). *Multiphoto Geometrically Constrained Matching*. Ph.d. thesis, Institute of Geodesy and Photogrammetry. ETH Zurich. Switzerland. 95, 198
- Baltsavias, E. P. (1999). *Matching Verfahren und automatische DTM Generierung* [Lehrmaterialien]. Zürich, CH: ETH, Eidgenössische Technische Hochschule Zürich, Institut für Geodäsie und Photogrammetrie. 83, 91, 97, 137
- Baltsavias, E. P. & Stallmann, D. (1992). Metric information extraction from SPOT images and the role of polynomial mapping functions. In *International Archives of Photogrammetry and Remote Sensing* (S. 358-364). 89
- Barnard, S. T. & Fischler, M. A. (1982). Computational stereo. *Computing Surveys*, 14 (4), 553-572. 211, 233
- Barron, J. L., Fleet, D. J. & Beauchemin, S. S. (1994). Performance of optical flow techniques. *International Journal of Computer Vision*, 12 (1), 43-77. 91, 93
- Batavia, P., Pomerleau, D. & Thorpe, C. (1997). Overtaking vehicle detection using implicit optical flow. In *Proceedings of the IEEE Transportation Systems Conference* (S. 729-734). 91
- Beder, C. & Förstner, W. (2006). Direkte Bestimmung von Zylindern aus 3D-Punkten ohne Nutzung von Oberflächennormalen. In T. Luhmann & C. Müller (Hrsg.), *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik* (S. 206-213). Herbert Wichmann Verlag. 49
- Beder, C. & Koch, R. (2008). Calibration of focal length and 3D pose based on the reflectance and depth image of a planar object. *International Journal of Intelligent Systems Technologies and Applications*, 5 (3/4), 285-294. 128, 132
- Bethmann, F. & Luhmann, T. (2010). Least-squares matching with advanced geometric transformation models. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII., S. 86-91). Newcastle upon Tyne, UK. 88
- Beuter, N., Swadzba, A., Kummert, F. & Wachsmuth, S. (2010). Using articulated scene models for dynamic 3D scene analysis in vista spaces. *3D Research*, 1, 20:1-20:13. 189
- Beyeler, D. (2006). *Single image 3d object tracking*. Diplomarbeit, Hochschule für Technik und Informatik, Software-Schule Schweiz. 186
- BGIA. (2006, März). *Einsatz des Messsystems CUELA zur Erfassung und Bewertung physischer Arbeitsbelastungen, Information des Berufsgenossenschaftlichen Instituts für Arbeitsschutz* (Information). Sankt Augustin, D: Berufsgenossenschaftliches Institut für Arbeitsschutz. 176
- Böhm, J. & Pattinson, T. (2010). Accuracy of exterior orientation for a range camera. In J. P. Mills, D. M. Barber, P. Miller & I. Newton (Hrsg.), *Proceedings of the ISPRS Commission V Mid-Term Symposium 'Close Range Image Measurement Techniques'* (Bd. XXXVIII,

- S. 103-108). 117, 122, 129, 132, 163, 165, 168
- Björck Åke. (1996). *Numerical Methods for Least Squares Problems*. Society for Industrial Mathematics. (Illustrated edition) 39
- Black, M., Yacoob, Y. & Ju, S. (1997). Recognizing human motion using parameterized models of optical flow. In Shah & Jain (Hrsg.), (Kap. 11). Kluwer Academic Publishing. 91
- Blanc, N. (2001). CCD versus CMOS - has CCD imaging come to an end. In D. Fritsch & R. Spiller (Hrsg.), *Photogrammetric Week '01*. Heidelberg, D: Herbert Wichmann Verlag. 106
- Böhme, M., Haker, M., Riemer, K., Martinetz, T. & Barth, E. (2009). Face detection using a time-of-flight camera. In *Dynamic 3D Imaging – Workshop in Conjunction with DAGM* (Bd. 5742, S. 167-176). 113
- Boker, S. M., Covey, E. S., Tiberio, S. S. & Deboeck, P. R. (2005). Synchronization in dancing is not winner-takes-all: Ambiguity persists in spatiotemporal symmetry between dancers. In *Proceedings of the North American Association for Computational, Social, and Organizational Science*. 20
- Bozek, P. (2010). *Implementierung und Evaluierung von Verfahren zur Bewegungsanalyse in 3D-Bildsequenzen*. Unveröffentlichte Diplomarbeit, Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Dresden, D. 89, 96
- Breidt, M., Bühlhoff, H. H. & Curio, C. (2010). Face models from noisy 3d cameras. In *ACM SIGGRAPH ASIA 2010 Sketches* (S. 12:1-12:2). New York, NY, USA: ACM. 188, 190
- Breiner, T. C. (2005). *Dreidimensionale virtuelle Organismen*. Dissertation, Johann Wolfgang Goethe Universität, Fachbereich Informatik und Mathematik, Frankfurt am Main. 68, 176, 182
- Breuer, P., Eckes, C. & Muller, S. (2007). Hand gesture recognition with a novel IR time-of-flight range camera: A pilot study. In *Computer Vision/Computer Graphics Collaboration Techniques: Third International Conference, MIRAGE 2007* (S. 247-260). 113, 188
- Brown, D. C. (1971). Close-range camera calibration. *Photogrammetric Engineering*, 37 (8), 855-866. 55
- Brown, D. C. (1976). The bundle adjustment - progress and prospects. In *International Archives of Photogrammetry* (Bd. 21). 60
- Büttgen, B., Oggier, T., Lehmann, M., Kaufmann, R. & Lustenberger, F. (2005). CCD/CMOS lock-in pixel for range imaging: Challenges, limitations and state-of-the-art. In H. Ingensand & T. Kahlmann (Hrsg.), *Range Imaging Research Day* (Bd. 1). Zürich, CH. 110
- Buehler, C., Matusik, W. & McMillan, L. (1999, March). *Creating and Rendering Image-Based Visual Hulls* (Technical Report Nr. 780). MIT LCS. 72
- Buhmann, E., Bachhuber, R. & Schaller, J. (Hrsg.). (1996). *ArcView: GIS-Arbeitsbuch*. Heidelberg: Wichmann. 69
- Burkhart, S. (2007). *Automatische Messung kreisförmiger Zielmarken*. Unveröffentlichte Studienarbeit, Technische Universität Dresden, Institut für Photogrammetrie und

Fernerkundung. 137

- Buxbaum, B. (2002). *Optische Laufzeitentfernungsmessung und CDMA auf Basis der PMD-Technologie mittels phasenvariabler PN-Modulation* (Bd. 17). Shaker Verlag. 110
- Canesta. (2010). Data sheet for the Canesta DP200 electronic perception technology development platform [Software-Handbuch]. Sunnyvale, CA, USA. Verfügbar unter <http://www.canesta.com/> 103, 104
- Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17 (8), 790-799. 189
- Chiabrando, F., Chiabrando, R., Piatti, D. & Rinaudo, F. (2009). Sensors for 3D imaging: Metric evaluation and calibration of a CCD/CMOS time-of-flight camera. *Sensors*, 9 (12), 80-96. 129
- Chiabrando, F., Piatta, D. & Rinaudo, F. (2010). Integration of ToF camera and multi-image matching approach for cultural heritage survey. In J. P. Mills, D. M. Barber, P. Miller & I. Newton (Hrsg.), *Proceedings of the ISPRS Commission V Mid-Term Symposium 'Close Range Image Measurement Techniques'* (Bde. Vol XXXVIII, Part 5, S. 143-148). 113
- Chum, O., Matas, J. & Kittler, J. (2003). Locally optimized RANSAC. In *DAGM-Symposium* (S. 236-243). 51
- D'Apuzzo, N. (2003). *Surface measurement and tracking of human body parts from multi station video sequences*. Dissertation, ETH Zurich, Institute of Geodesy and Photogrammetry, Zurich, CH. 69, 98, 180, 181
- D'Apuzzo, N., Plänklers, R. & Fua, P. (2000). Least squares matching tracking algorithm for human body modeling. In *International archives of photogrammetry and remote sensing* (Bd. 33, S. 714-720). 94
- Dimitrijevic, M., Lepetit, V. & Fua, P. (2005). Human body pose recognition using spatio-temporal templates. In *ICCV workshop on Modeling People and Human Interaction, Beijing, China*. 178, 179
- Diraco, G., Leone, A. & Siciliano, P. (2010). An active vision system for fall detection and posture recognition in elderly healthcare. In *Proceedings of the Conference on Design, Automation and Test in Europe* (S. 1536-1541). 3001 Leuven, Belgium, Belgium: European Design and Automation Association. 188, 190
- Doherty-Sneddon, G. & Phelps, F. G. (2007). Teachers responses to childrens eye gaze. *Educational Psychology*, 27 (1), 97-109. 176
- Dold, C. & Brenner, C. (2006). Registration of terrestrial laser scanning data using planar patches and image data. In H.-G. Maas & D. Schneider (Hrsg.), *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. Volume XXXVI, S. 78-83). 98
- Dracos, T., Merkel, G., Rys, F., Rys, P., Maas, H.-G. & Grün, A. (1995). Velocity field measurement using laser induced fluorescence tomography. In *Proceedings of the International Workshop PEV, Fukui 95*. Japan. 97

- Dörfler, M., Dittrich, G. & Schneider, K. (2001). *Konflikte unter Kindern – Ein Kinderspiel für Erwachsene?* [Arbeitsmaterialien für Erzieherinnen]. Beltz. 21, 226
- Droeschel, D., Holz, D. & Behnke, S. (2010, June). Probabilistic phase unwrapping for time-of-flight cameras. In *Proceedings of the joint conference of the 41st International Symposium on Robotics (ISR 2010) and the 6th German Conference on Robotics (ROBOTIK 2010)* (S. 318–324). Munich, Germany. 109
- Du, H., Oggier, T., Lustenberger, F. & Charbon, E. (2005). A virtual keyboard based on true-3D optical ranging. In *British Machine Vision Conference 2005* (Bd. 1, S. 220-229). 113, 123
- Ebner, H. (1976). Bildmessung und Luftbildwesen. *Bildmessung und Luftbildwesen*, 44, 128-139. 197
- Ekman, P., Friesen, W. V. & Hager, J. C. (1978). *The facial action coding system* (1. Aufl.). Palo Alto, CA, USA: Consulting Psychologists Press. 20
- El-Hakim, S. (1986). A real-time system for object measurement with CCD cameras. *IAPRS*, 26. 55
- Fanpop, I. (2011). Zugriff am 03.02.2011 auf <http://www.fanpop.com/spots/avatar/links/11559236> 78
- Fischler, M. A. & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24 (6), 381-395. 48, 49
- Foresti, G. L., Micheloni, C., Snidaro, L., Remagnino, P. & Ellis, T. (2005). Active video-based surveillance system. 81
- Frank, H., Wellerdick-Wojtasik, N., Hagebecker, B., Novak, G. & Mahlke, S. (2006). Throwing objects - a bio-inspired approach for the transportation of parts. In *Proceedings of the 2006 IEEE international conference on robotics and biomimetics (robio)* (S. 91-96). Los Alamitos, CA, USA: IEEE Computer Society. 112
- Fraser, C. (1997). Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52 (4), 149-159. 127
- Freedman, B., Shpunt, A. & Arieli, Y. (2010). *Distance-varying Illumination and Imaging Techniques for Depth Mapping* (Nr. US 2010/0290698 A1). Tel Aviv, IL. 242
- Frey, S., Hirsbrunner, H.-P., Pool, J. & Daw, W. (1981). Das Berner System zur Untersuchung nonverbaler Interaktion: I. Die Erhebung des Rohdatenprotokolls. In P. Winkler (Hrsg.), *Methoden der Analyse von Face-to-Face Situationen* (S. 203-236). Stuttgart, D: Metzler. 20
- Frischholz, R. (1999). Motion Tracking. In B. Jähne, H. Haußecker & P. Geißler (Hrsg.), *Handbook of Computer Vision and Applications - Systems and Applications* (Bd. 3, S. 329-344). Academic Press. 176
- Förstner, W. (1979). Ein Verfahren zur Schätzung von Varianz- und Kovarianzkomponenten. *Allgemeine Vermessungsnachrichten*, 86, 41-49. 45, 48, 156, 209
- Förstner, W. (1986). A feature based correspondence algorithm for image matching. In *Archives of*

- Photogrammetry and Remote Sensing* (Bd. XXVI, S. 150-166). 94
- Förstner, W. & Gülch, E. (1987). A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *ISPRS Intercommission Workshop* (S. 281-305). Interlaken, CH. 85
- Fua, P., Herda, L., Plaenkers, R. & Boulic, R. (2000). Human shape and motion recovery using animation models. In *International Archives of Photogrammetry and Remote Sensing* (Bd. 33, S. 253-268). 180, 181
- Fuchs, S. & Hirzinger, G. (2008). Extrinsic and depth calibration of ToF-cameras. In *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008* (S. 1-6). 121, 122, 123, 129, 132
- Fuchs, S. & May, S. (2007). Calibration and registration for precise surface reconstruction with ToF cameras. In *Proceedings of the Dynamic 3D Imaging Workshop*. Heidelberg, D. 129
- Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition* (2. Aufl.). San Diego, CA, USA: Academic Press Professional, Inc. 228
- Gall, J., Rosenhahn, B., Gehrig, S. & Seidel, H.-P. (2008). Model-based motion capture for crash test video analysis. In *30th Annual Symposium of the German Association for Pattern Recognition (DAGM'08)* (S. 92-101). Springer. 78
- Gallo, L., Placitelli, A. P. & Ciampi, M. (2011). Controller-free exploration of medical image data: Experiencing the kinect. In *Proceedings of the 24th IEEE International Symposium on Computer-Based Medical Systems* (S. 1-6). Bristol, UK. 243, 244
- Ganapathi, V., Plagemann, C., Koller, D. & Thrun, S. (2010). Real time motion capture using a single time-of-flight camera. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 755-762. 186, 187
- Garbe, C. S., Roetmann, K., Beushausen, V. & Jähne, B. (2008). An optical flow MTV based technique for measuring microfluidic flow in the presence of diffusion and Taylor dispersion. *Experiments in Fluids*, 44 (3), 439-450. 91
- Gauß, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Hamburg. 35
- Gavrila, D. M. (1996). *Vision-based 3-D Tracking of Humans in Action*. Ph.d. thesis, Department of Computer Science, University of Maryland, College Park, Maryland, USA. 180
- Godding, R. (1993). Ein photogrammetrisches System zur Überprüfung und Kalibrierung digitaler Bildaufnahmesysteme. *Zeitschrift für Photogrammetrie und Fernerkundung*, 2, 82-90. 62
- Göktürk, S. B. & Rafii, A. (2005). An occupant classification system eigen shapes or knowledge-based features. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops* (S. 57). Washington, DC, USA: IEEE Computer Society. 112, 186
- Göktürk, S. B. & Tomasi, C. (2004). 3D head tracking based on recognition and interpolation using a time-of-flight depth sensor. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2004 CVPR 2004* (Bd. II, S. 211-217). 112



- Gordon, B. (2008). *Zur Bestimmung von Messunsicherheiten terrestrischer Laserscanner*. Unveröffentlichte Dissertation, TU Darmstadt. 160
- Grammer, K., Keki, V., Striebel, B., Atzmüller, M. & Fink, B. (2003). Bodies in motion: A window to the soul. In E. Voland & K. Grammer (Hrsg.), *Evolutionary Aesthetics* (S. 295-324). Heidelberg, D: Springer. 20
- Grest, D., Krüger, V. & Koch, R. (2007). Single view motion tracking by depth and silhouette information. In *SCIA* (S. 719-729). 187, 188
- Grün, A. (1985). Adaptive least squares correlation - a powerful image matching technique. In *South African Journal of Photogrammetry, Remote Sensing and Cartography* (Bd. 14, S. 175-187). 94, 98
- Grün, A. (1986). *Photogrammetrische Punktbestimmung mit der Bündelmethode* (Mitteilungen Nr. 40). Zürich, CH: Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich. 197
- Grün, A. & Akca, D. (2004). Least squares 3d surface matching. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXIV). 98, 196
- Grynszpan, O., Martin, J.-C. & Oudin, N. (2003). On the annotation of gestures in multimodal autistic behaviour. In *The 5th International Workshop on Gesture and Sign Language based Human-Computer Interaction*. Genova, I. 176
- Gudmundsson, S. A., Larsen, R., Aanaes, H., Pardas, M. & Casas, J. (2008). ToF imaging in smart room environments towards improved people tracking. In *Computer Vision and Pattern Recognition Workshops, 2008, CVPR Workshop on Time of Flight Camera based Computer Vision* (S. 1-6). 188
- Gudmundsson, S. A., Larsen, R. & Ersbøll, B. K. (2007). Robust pose estimation using the swissranger SR-3000 camera. In *SCIA'07: Proceedings of the 15th Scandinavian conference on Image analysis* (S. 968-975). Berlin, Heidelberg: Springer-Verlag. 112, 119
- Gudmundsson, S. A., Sveinsson, J. R., Pardas, M., Aanaes, H. & Larsen, R. (2010). Model-based hand gesture tracking in tof image sequences. In *Proceedings of the 6th International Conference on Articulated Motion and Deformable Objects* (S. 118-127). Berlin, Heidelberg: Springer-Verlag. 186
- Gut, O. (2004). *Untersuchungen des 3D-Sensors SwissRanger*. Diplomarbeit, ETH Zürich - Eidgenössische Technische Hochschule Zürich, Institut für Geodäsie und Photogrammetrie, Professur für Geodätische Messtechnik. 117
- Hampel, U. (2008). *Photogrammetrische Erfassung der Verformungs- und Rissentwicklung bei baumechanischen Untersuchungen*. Dissertation, Technische Universität Dresden, Fakultät Bauingenieurwesen, Dresden. 69
- Hasanlou, M. & Saradjian, M. (2006). Water motion analysis in SST images using least squares methods. In N. Kerle & A. K. Skidmore (Hrsg.), *Proceedings of the ISPRS mid-term Symposium* (Bd. XXXVI). Enschede, NL. 89
- Haußecker, H. & Spies, H. (1999). Motion. In B. Jähne, H. Haußecker & P. Geißler (Hrsg.),



- Handbook of Computer Vision and Applications - Signal Processing and Pattern Recognition* (Bd. 2, S. 309-396). Academic Press. 79, 83, 91
- Heckenkamp, C. (2008). Das magische Auge - Grundlagen der Bildverarbeitung: Das PMD-Prinzip. *INSPECT, 1*, 25-28. 105
- Heinol, H. G. (2001). *Untersuchung und Entwicklung von modulationslaufzeitbasierten 3D-Sichtsystemen*. Dissertation, Department of Electrical Engineering and Computer Science. 106
- Heinol, H. G. & Schwarte, R. (1999). Photomischdetektor erfaßt 3D-Bilder: Neues optisches Bauelement vereinigt Detektion und Mischung. *Elektronik, WEKA Fachzeitschriften-Verlag München, 48* (12). 105
- Helava, U. V. (1988). Object space least squares correlation. *Int. Arch. Photogramm. Remote Sensing, 27* (B2), 297-302. 94
- Hempel, M. (2007). *Validierung der Genauigkeit und des Einsatzpotentials einer distanzmessenden Kamera*. Unveröffentlichte Diplomarbeit, Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Dresden, D. 117, 119, 154, 165, 198, 206
- Hempel, R. & Westfeld, P. (2009). Statistical modeling of interpersonal distance with range imaging data. In A. Esposito, A. Hussain, M. Marinaro & R. Martone (Hrsg.), *Multimodal Signals: Cognitive and Algorithmic Issues. COST Action 2102 and euCognition International School Vietri sul Mare, Italy, April 21-26, 2008 Revised Selected and Invited Papers* (Bd. 5398/2009, S. 137-144). Springer Berlin / Heidelberg. 113, 154, 206, 226, 227
- Henry, P., Krainin, M., Herbst, E., Ren, X. & Fox, D. (2010). RGB-D Mapping: Using depth cameras for dense 3D modeling of indoor environments. In *Proceedings of the 12th International Symposium on Experimental Robotics*. 243, 244
- Hilsmann, A. & Eisert, P. (2008). Optical flow based tracking and retexturing of garments. In *Proc. International Conference on Image Processing ICIP* (S. 845-848). San Diego, California, USA. 89
- Hofmann, S. (2008). *Konzeption und Aufbau eines optischen 3D Messsystems zur Analyse orthopädisch relevanter Bewegungsparameter*. Unveröffentlichte Diplomarbeit, Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Dresden, D. 78
- Hogrefe, K., Weidinger, N., Lindner, K., Ziegler, W. & Goldenberg, G. (2011). *Speech replacing gestures: What are the underlying cognitive and neural processes?* [Projektseite]. München, D: Ludwig-Maximilians-Universität München. Verfügbar unter [http://www.ekn.mwn.de/index.php?option=com\\_content&view=article&id=9&Itemid=18&lang=en](http://www.ekn.mwn.de/index.php?option=com_content&view=article&id=9&Itemid=18&lang=en) 21
- Holte, M. B. & Moeslund, T. B. (2007). *Gesture Recognition using a Range Camera* (Technical Report Nr. CVMT-07-01 ISSN 1601-3646). Aalborg University, Denmark. 113
- Holte, M. B., Moeslund, T. B. & Fihl, P. (2010). View-invariant gesture recognition using 3d optical flow and harmonic motion context. *Comput. Vis. Image Underst., 114*, 1353-1361. 189, 190
- Horn, B. K. P. (1970). *Shape From Shading: A Method for Obtaining the Shape of a Smooth Opaque Object From One View* (Bericht). Cambridge, MA, USA. 69, 71

- Horn, B. K. P. & Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17, 185-203. 84, 92, 93
- Howe, N. R., Leventon, M. E. & Freeman, W. T. (1999). Bayesian reconstruction of 3d human motion from single-camera video. *Artificial Intelligence*, 1, 820-826. 178
- Hsu, S., Acharya, S., Rafii, A. & New, R. (2006). Performance of a time-of-flight range camera for intelligent vehicle safety applications. In *Advanced Microsystems for Automotive Applications* (S. 205-219). 112
- Huber, D., Kuffner, J. & Vandapel, N. (2004). *Real Time 3D Vision for Humanoid Robots*. 113
- Hussmann, S. & Edeler, T. (2010). Pseudo-four-phase-shift algorithm for performance enhancement of 3d-tof vision systems. *IEEE Transactions on Instrumentation and Measurement*, 59 (5), 1175-1181. 121
- Hussmann, S., Hermanski, A. & Edeler, T. (2010). Real-time motion suppression in tof range images. In *Instrumentation and Measurement Technology Conference (I2MTC), 2010 IEEE* (S. 697 - 701). Austin, TX, USA. 125
- Inition: Everything in 3D. (2011). *Vicon Peak - Vicon MX*. London, UK. Zugriff am 29.03.2011 auf [http://www.inition.co.uk/inition/product\\_mocaptrack\\_vicon\\_motioncapture.htm](http://www.inition.co.uk/inition/product_mocaptrack_vicon_motioncapture.htm) 183
- Isard, M. & Blake, A. (1998). Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29 (1), 5-28. 186
- Jaakkola, A., Kaasalainen, S., Hyyppä, J., Akujärvi, A. & Niittymäki, H. (2008). Radiometric calibration of intensity images of SwissRanger SR-3000 range camera. *The Photogrammetric Journal of Finland*, 21 (1), 16-25. 124
- Jensen, R. R., Paulsen, R. R. & Larsen, R. (2009). Analysis of gait using a treadmill and a time-of-flight camera. In *Proceedings of the DAGM 2009 Workshop on Dynamic 3D Imaging* (S. 154-166). Berlin, Heidelberg: Springer-Verlag. 191, 194
- Jäger, R., Müller, T., Saler, H. & Schwäble, R. (2005). *Klassische und robuste Ausgleichungsverfahren*. Heidelberg: Herbert Wichmann Verlag. 43
- Jähne, B. (1999). Spatial and fourier domain. In B. Jähne, H. Haußecker & P. Geißler (Hrsg.), *Handbook of Computer Vision and Applications - Signal Processing and Pattern Recognition* (Bd. 2, S. 35-66). Academic Press. 96
- Jähne, B. (2002). *Digitale Bildverarbeitung* (5. Aufl.). Berlin, Heidelberg, BRD: Springer-Verlag. 80, 83, 84, 87, 91, 98, 120
- Jähne, B., Klar, M. & Jehle, M. (2007). Handbook of Experimental Fluid Mechanics. In C. Tropea, J. Foss & A. Yarin (Hrsg.), (S. 1437-1491). Springer. 84
- Joeckel, R. & Stober, M. (1999). *Elektronische Entfernungsmessung* (4. Aufl.). Konrad Wittwer. (ISBN 3879192669) 72, 73, 74, 121, 122
- Justen, D. (2001). *Untersuchung eines neuartigen 2D- gestützten 3D-PMD-Bildverarbeitungssysteme*. Dissertation, Department of Electrical Engineering and Computer Science, Siegen. 106
- Jutzi, B. (2009). Investigations on ambiguity unwrapping of range images. In F. Bretar,

- M. Pierrot-Deseilligny & G. Vosselman (Hrsg.), *International archives of photogrammetry and remote sensing* (Bd. 38, S. 265). 109
- Kaasalainen, S., Kukko, A., Lindroos, T., Litkey, P., Kaartinen, H., Hyypä, J. et al. (2008). Brightness measurements and calibration with airborne and terrestrial laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, 46 (2), 528-534. 124
- Kadous, M. W., Sheh, R. K.-M. & Sammut, C. (2006). Controlling heterogeneous semi-autonomous rescue robot teams. In *Proceedings IEEE Conference on Systems, Man and Cybernetics*. Taiwan. 112
- Kahlmann, T. (2007). *Range Imaging Metrology – Investigation, Calibration and Development*. Dissertation, nr. 17392, Eidgenössische Technische Hochschule ETH Zürich. 106, 107, 108, 110, 117, 118, 121, 124, 132, 163
- Kahlmann, T. & Ingensand, H. (2004). Calibration and improvements of the high-resolution range-imaging camera SwissRanger. In J.-A. Beraldin, S. F. El-Hakim, A. Grün & J. S. Walton (Hrsg.), *Videometrics VIII* (Bd. 5665, S. 144-155). 128
- Kahlmann, T. & Ingensand, H. (2006). Calibration of the fast range imaging camera SwissRanger for use in the surveillance of the environment. In G. W. Kamerman, D. V. Willetts & O. K. Steinvall (Hrsg.), *Electro-Optical Remote Sensing II. Proceedings of the SPIE* (Bd. 6396, S. 5). 113, 128, 186
- Kahlmann, T. & Ingensand, H. (2007a). Increased accuracy of 3D range imaging camera by means of calibration. In *Optical 3-D Measurement Techniques VIII*. 124
- Kahlmann, T. & Ingensand, H. (2007b). Range Imaging Metrologie: Einführung, Untersuchungen und Weiterentwicklungen. *Allgemeine Vermessungs-Nachrichten*, 11-12, 284-393. 119
- Kahlmann, T., Remondino, F. & Guillaume, S. (2007). Range imaging technology: new developments and applications for people identification and tracking. In A. Beraldin, F. Remondino & M. R. Shortis (Hrsg.), *Proceedings of Videometrics IX - SPIE-IS&T Electronic Imaging* (Bd. 6491). 113, 121, 186, 187
- Kahlmann, T., Remondino, F. & Ingensand, H. (2006). Calibration for increased accuracy of the range imaging camera swissrangertm. In H.-G. Maas & D. Schneider (Hrsg.), *ISPRS Image Engineering and Vision Metrology* (Bde. XXXVI, 5, S. 136-141). (ISSN 1682-1750) 110, 113, 122, 128
- Kahmen, H. (1997). *Vermessungskunde* (Bd. 19. Auflage). Berlin, New York: Walter de Gruyter. 40, 75
- Karel, W. (2008). Integrated range camera calibration using image sequences from hand-held operation. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVII, S. 945-951). 129
- Karel, W., Ghuffar, S. & Pfeifer, N. (2010). Quantifying the distortion of distance observations caused by scattering in time-of-flight range cameras. In J. P. Mills, D. M. Barber, P. Miller & I. Newton (Hrsg.), *Proceedings of the ISPRS Commission V Mid-Term Symposium 'Close Range Image Measurement Techniques'* (Bd. XXXVIII, S. 316-321). 118, 119, 120

- Karel, W. & Pfeifer, N. (2009). Range camera calibration based on image sequences and dense, comprehensive error statistics. In A. Beraldin, G. S. Cheok, M. McCarthy & U. Neuschaefer-Rube (Hrsg.), *SPIE Proceedings on Electronic Imaging / 3D Imaging Metrology* (Bd. 7239). San José, CA, USA. 124
- Kavli, T., Kirkhus, T., Thielemann, J. T. & Borys, J. (2008). Modelling and compensating measurement errors caused by scattering in time-of-flight cameras. In *Two- and Three-Dimensional Methods for Inspection and Metrology VI, Proceedings of SPIE, the International Society for Optical Engineering*. 119, 120
- Kawaguchi, H., Murakami, B. & Kawai, M. (2010). Behavioral characteristics of children with high functioning pervasive developmental disorders during a game. *Journal of Epidemiology*, 20, 490-497. 20, 22
- Kempton, W. (1980). The rhythmic basis of interactional micro-synchrony. In J. A. Fisherman & M. R. Key (Hrsg.), *The Relationship of Verbal and Nonverbal Communication*. Den Haag, NL: Mouton Publishers. 20
- Khoshelham, K. (2011). Accuracy analysis of kinect depth data. In D. D. Lichti & A. F. Habib (Hrsg.), *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII). Calgary, Canada. 243
- Kim, Y. M., Chan, D., Theobalt, C. & Thrun, S. (2008). Design and calibration of a multi-view tof sensor fusion system. In *IEEE CVPR Workshop on Time-of-flight Computer Vision* (S. 1-7). 122
- KinectHacks. (2011). *Kinect Hacks - Supporting the Kinect Hacking news and community*. Zugriff am 04.10.2011 auf <http://kinecthacks.net/> 245
- Kitzhofer, J., Westfeld, P., Pust, O., Nonn, T., Maas, H.-G. & Brücker, C. (2010). Estimation of 3D deformation and rotation rate tensor from volumetric particle data via 3D least squares matching. In *Proceedings of the 15th Int Symp on Applications of Laser Techniques to Fluid Mechanics*. Lisbon, Portugal. 97, 98
- Klein, B. (2001). *Untersuchungen zur Feldprüfung geodätischer Instrumente mittels Varianzkomponentenschätzung*. Diplomarbeit, Technische Universität Darmstadt. 45
- Klette, R., Kozera, R. & Schlüns, K. (1999). Reflectance-based shape recovery. In B. Jähne, H. Haußecker & P. Geißler (Hrsg.), *Handbook of Computer Vision and Applications - Signal Processing and Pattern Recognition* (Bd. 2, S. 531-590). Academic Press. 70, 71
- Knoop, S., Vacek, S. & Dillmann, R. (2006). Sensor fusion for 3d human body tracking with an articulated 3d body model. *Proceedings 2006 IEEE International Conference on Robotics and Automation*, 1686-1691. 188
- Koch, K.-R. (2004). *Parameterschätzung und Hypothesentests in linearen Modellen* (4. Aufl.). Bonn: Ferd. Dümmlers Verlag. 34, 45, 47, 48, 156, 209
- Kohoutek, T. K. (2008). Analysis and processing the 3D-range-image-data for robot monitoring. *Geodesy and Cartography - Research Journal of Vilnius Gediminas Technical University*, 34 (3), 92-96. 112, 186

- Kohoutek, T. K. (2009). Multi-user vision interface based on range imaging. In *Visual Information Processing XVIII, Proceedings of SPIE* (Bd. 7341, S. 73410). 113
- Kohoutek, T. K., Mautz, R. & A., D. (2010). Real-time indoor positioning using range imaging sensors. In *Real-Time Image and Video Processing, Proceedings of SPIE* (Bd. 7724). 114
- Kohoutek, T. K. & Nitsche, M. (2010). Assessment of flow resistance of steep mountain rivers measuring streambed topography by range imaging technology. In J. P. Mills, D. M. Barber, P. E. Miller & I. Newton (Hrsg.), *Proceedings of the ISPRS Commission V Mid-Term Symposium 'Close Range Image Measurement Techniques'* (Bd. XXXVIII, S. 355-360). 114
- Koller-Meier, E. B. (2000). *Extending the Condensation Algorithm for Tracking Multiple Objects in Range Image Sequences*. Unveröffentlichte Dissertation, Technische Wissenschaften ETH Zürich, Nr. 13548, Zürich, CH. 186
- Kommer, S. & Biermann, R. (2005). Video(technik) in der erziehungswissenschaftlichen Forschung. *MedienPädagogik - Zeitschrift für Theorie und Praxis der Medienbildung. Pädagogische Hochschule Zürich..* 176
- Konolige, K. & Mihelich, P. (2010). *Technical aspects of the kinect device and its calibration* (Bericht). Verfügbar unter [http://www.ros.org/wiki/kinect\\_calibration/technical](http://www.ros.org/wiki/kinect_calibration/technical) 242
- Kraus, K. (1996). *Photogrammetrie - Verfeinerte Methoden und Anwendungen* (Bd. 2). Bonn: Ferd. Dummlers Verlag. 44, 99, 197
- Krzystek, P. (1991). Fully automatic measurement of digital elevation models with MATCH-T. In *Photogrammetric Week 1991 - Schriftenreihe des Instituts für Photogrammetrie* (S. 203-214). Stuttgart, BRD. 85
- Kubacki, J. & Pfeifer, K. (2005). Using range imaging sensors with color imaging sensors in cognitive robot companions: A new and simple calibration technique based on particle swarm optimization. In H. Ingensand & T. Kahlmann (Hrsg.), *Range Imaging Research Day* (Bd. 1). Zurich. 113
- Kubik, K. (1967). Schätzung der Gewichte der Fehlergleichungen beim Ausgleichungsproblem nach vermittelnden Beobachtungen. *Zeitschrift für Vermessungswesen*, 92, 173-178. 45
- Kutulakos, K. N. & Seitz, S. M. (2000). A theory of shape by space carving. *International Journal of Computer Vision*, 38 (3), 199-218. 71
- LaFrance, M. & Broadbent, M. (1976). Group rapport: Posture sharing as a nonverbal indicator. *Group Organization Management*, 1 (3), 328. 175
- Lahamy, H. & Lichti, D. D. (2011). Evaluation of real-time hand motion tracking using a range camera and the mean-shift algorithm. In D. D. Lichti & A. F. Habib (Hrsg.), *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII). Calgary, Canada. 189
- Lange, R. (2000). *3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology*. Phd thesis, Department of Electrical Engineering and Computer Science at University of Siegen. 106, 108, 110, 117, 120, 122, 123, 160
- Lange, R., Seitz, P. & Schwarte, R. (1999). Time-of-flight Entfernungskamera in CMOS/CCD-



- Technik mit pixelintegrierten optischen Lock-in-Verstärkern. In *Fachtagung optische formerfassung* (Bd. 70, S. 271-280). Stuttgart, D.. 103
- Laurentini, A. (1994). The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16 (2), 150–162. 71
- Laurentini, A. (1997). How many 2D silhouettes does it takes to reconstruct a 3D object? *Computer Vision and Image Understanding*, 67 (1), 81-87. 71
- Lausberg, H. & Kryger, M. (2011). Gestisches Verhalten als Indikator therapeutischer Prozesse in der verbalen Psychotherapie: Zur Funktion der Selbstberührungen und zur Repräsentation von Objektbeziehungen in gestischen Darstellungen. *Psychotherapie-Wissenschaft*, 1 (1), 41-55. 20, 21
- Lausberg, H. & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, 41 (3), 841-849. 176, 230, 231
- Lausberg, H., Zaidel, E., Cruz, R. F. & Ptito, A. (2007). Speech-independent production of communicative gestures: Evidence from patients with complete callosal disconnection. *Neuropsychologia*, 45 (13), 3092-3104. 19, 176, 228
- Lawson, C. L. & Hanson, R. J. (1995). *Solving Least Squares Problems (Classics in Applied Mathematics)*. Society for Industrial Mathematics. 39
- Li, R. (1992). Reconstruction of seafloor surface models by shape from shading. In *ISPRS Archives* (Bd. XXIX, S. 780-784). 70
- Liang, T. & Heipke, C. (1996, January). Automatic relative orientation of aerial images. *Photogrammetric engineering and remote sensing*, 62 (1), 47-55. 89
- Lichti, D. D. (2008). Self-calibration of a 3d range camera. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. 37, S. 927-932). 130
- Lichti, D. D., Kim, C. & Jamtsh, S. (2010). An integrated bundle adjustment approach to range camera geometric self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65 (4), 360-368. 117, 122, 130, 132, 163, 165, 168
- Lindner, M. & Kolb, A. (2006). Lateral and depth calibration of PMD-distance sensors. In *International Symposium on Visual Computing (ISVC06)* (Bd. 2, S. 524-533). Springer. 122, 128
- Lindner, M. & Kolb, A. (2007). Calibration of the intensity-related distance error of the pmd tof-camera. In *SPIE: Intelligent Robots and Computer Vision XXV* (Bd. 6764, S. 6764-35). 124, 128
- Lindner, M. & Kolb, A. (2009). Compensation of Motion Artifacts for Time-of-Flight Cameras. In *Proc. Dynamic 3D Imaging* (S. 16-27). Springer. 125
- Lottner, O., Sluiter, A., K., H. & Weihs, W. (2007). Movement artefacts in range images of time-of-flight cameras. In *International Symposium on Signals, Circuits & Systems - ISSCS 2007* (Bd. 1, S. 117-120). 125
- Luan, X. (2001). *Experimental investigation of photonic mixer device and development of TOF 3D*

- ranging systems based on pmd technology*. Dissertation, Department of Electrical Engineering and Computer Science, Universität Siegen, Siegen, D. 117, 121
- Lucas, B. D. & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the 1981 DARPA Image Understanding Workshop* (S. 674-679). 84, 92
- Luhmann, T. (2003). *Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen* (2. Aufl.). Wichmann. 32, 34, 54, 55, 57, 59, 60, 69, 75, 84, 85, 91
- Maas, H.-G. (1992a). Complexity analysis for the determination of image correspondences in dense spatial target fields. In *International Archives of Photogrammetry and Remote Sensing* (Bd. 29, S. 102-107). 89
- Maas, H.-G. (1992b). Schnelle Oberflächenvermessung mittels optischer Sensoren. In H. Matthias & A. Grün (Hrsg.), *XI. Internationaler Kurs für Ingenieurvermessung* (S. II 3/1 - II 3/12). Zürich, CH. 66
- Maas, H.-G. (1993). Bestimmung dreidimensionaler Geschwindigkeitsfelder aus Strömungstomographie-Sequenzen. In *DAGM-Symposium* (S. 359-366). 97
- Maas, H.-G. (1996a). Adaptive least-squares-matching for accurate tracking of patterns in voxel data. In T. Dracos (Hrsg.), *Three-dimensional Velocimetry and Vorticity Measuring and Image Analysis Techniques* (S. 265-276). Dordrecht: Kluwer Academic Publishers. 97
- Maas, H.-G. (1996b). Automatic dem generation by multi-image feature based matching. In *International Archives of Photogrammetry and Remote Sensing* (Bd. XXXI, S. 484-489). 85
- Maas, H.-G. (1997). *Mehrbildtechniken in der digitalen Photogrammetrie*. Habilitationsschrift, Eidg. Technischen Hochschule Zürich, Institut für Geodäsie und Photogrammetrie, Zürich, CH. 55, 59, 70, 80, 87, 115, 182
- Maas, H.-G. (2000). Least-squares matching with airborne laserscanning data in a tin structure. *International Archives of Photogrammetry and Remote Sensing*, 33 (3a), 548-555. 98, 99, 192
- Maas, H.-G. (2001). On the use of pulse reflectance data for laserscanner strip adjustment. In *ISPRS Workshop on Land Surface Mapping and Reconstruction using Laser Altimetry* (Bd. XXXIV). Annapolis/Maryland, USA. 98, 192, 196
- Maas, H.-G. (2005a). Akquisition von 3D-GIS Daten durch Flugzeuglaserscanning. *Allgemeine Vermessungs-Nachrichten*, 55 (1), 3-11. 75
- Maas, H.-G. (2005b). Werkzeuge und Anwendungen der photogrammetrischen 3D-Bewegungsanalyse. In T. Luhmann (Hrsg.), *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik (Beiträge Oldenburger 3D-Tage 2005)* (S. 2-12). Verlag Herbert Wichmann. 180
- Maas, H.-G., Dietrich, D., Schwalbe, E., Bäsler, M. & Westfeld, P. (2005). Photogrammetrische Bestimmung räumlich-zeitlich aufgelöster Bewegungsfelder am Jakobshavn Isbræ Gletscher in Grönland. In *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation* (Bd. 13). 88
- Maas, H.-G., Grün, A. & Papantoniou, D. (1993). Particle tracking in threedimensional turbulent



- flows. In *Experiments in Fluids* (Bd. 15, S. 133-146). 180
- Maas, H.-G., Schneider, D., Schwalbe, E., Casassa, G. & Wendt, A. (2010). Photogrammetric techniques for the determination of spatio-temporal velocity fields at Glacier San Rafael in the Northern Patagonian Icefield. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38. 78
- Maas, H.-G., Stefanidis, A. & Grün, A. (1994). From pixels to voxels - tracking volume elements in sequences of 3-d digital images. In *International Archives of Photogrammetry and Remote Sensing* (Bd. 30, S. 539-546). 97
- May, S., Droschel, D., Holz, D., Fuchs, S., Malis, E., Nüchter, A. et al. (2009). Three-dimensional mapping with time-of-flight cameras. *Journal of Field Robotics, Special Issue on Three-Dimensional Mapping*, 26 (11-12), 934-965. 112, 121
- McGlone, J. C., Mikhail, E. M. & Bethel, J. (Hrsg.). (2004). *Manual of Photogrammetry* (5. Aufl.). Bethesda, Maryland, USA: American Society for Photogrammetry and Remote Sensing. 83, 91, 97
- Meers, S. & Ward, K. (2008). Head-pose tracking with a time-of-flight camera. In *Proceedings of the Australian Conference on Robotics and Automation*. Canberra, AUS. 189
- Meißner, T. (2010). *Integrierte Kalibrierung einer distanzmessenden Kamera*. Unveröffentlichte Diplomarbeit, Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Dresden, D. 159, 162
- MESA. (2010). SR4000 user manual [Software-Handbuch]. Zurich. Verfügbar unter <http://www.mesa-imaging.ch/> 103, 104, 106, 107, 118, 144
- Microsoft. (2011). *Xbox 360 - Official Site*. Zugriff am 04.10.2011 auf <http://www.xbox.com/> 244
- Mitchell, H. (2010). Simple shape-from-shading in the dental and medical measurement context. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII, S. 458-461). 70
- Miyagawa, R. & Kanade, T. (1997). CCD-based range-finding sensor. *IEEE Transactions on Electron Devices*, 44 (1), 1648-1652. 105
- Möller, T., Kraft, H., Frey, J., Albrecht, M. & Lange, R. (2005). Robust 3D measurement with PMD sensors. In H. Ingensand & T. Kahlmann (Hrsg.), *Range Imaging Research Day* (Bd. 1). Zurich, CH. 106
- Moeslund, T. B., Hilton, A. & Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Underst.*, 104, 90-126. 175
- Mollath, G. & Fatehi, D. (1994). Anwendung des Lichtschnittverfahrens zur Überwachung von Förderbändern. In *Optisches Messen von Länge und Gestalt. Erfahrungsaustausch zwischen Anwendern und Entwicklern. Tagungsband Düsseldorf, 1994 (GMA-Bericht 23)*. Düsseldorf. 68
- Motion Analysis Corporation. (2011). Zugriff am 03.02.2011 auf <http://www.motionanalysis.com/> 182, 183
- Möser, M., Müller, G., Schlemmer, H. & Werner, H. (2000). *Handbuch Ingenieurgeodäsie -*

- Grundlagen* (Bd. 3). Heidelberg: Herbert Wichmann Verlag. 43, 66
- Mulsoy, C. (2010). A flexible multi-media bundle approach. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII). Newcastle upon Tyne, UK. 60
- Mulsoy, C., Maas, H.-G., Westfeld, P. & Schulze, M. (2008). Triangulation methods for height profile measurements on instationary water surfaces. *Journal of Applied Geodesy*, 2, 21-29. 68
- Mure-Dubois, J. & Hügli, H. (2007). Real-time scattering compensation for time-of-flight camera. In *The 5th International Conference on Computer Vision Systems*. 119
- Netzsch, T. & Jähne, B. (1995). A high performance system for 3-dimensional particle tracking velocimetry in turbulent flow research using image sequences. In *International Archives of Photogrammetry and Remote Sensing* (Bd. 30). 180
- Nevatia, R. (1982). *Machine Perception*. Prentice Hall. 91
- Niclass, C., Desse, P.-A. & Charbon, E. (2005). Arrays of single photon avalanche diodes in CMOS technology: Picosecond timing resolution for range imaging. In H. Ingensand & T. Kahlmann (Hrsg.), *Range Imaging Research Day* (Bd. 1). Zürich, CH. 103
- Niemeier, W. (2002). *Ausgleichsrechnung - Eine Einführung für Studierende und Praktiker des Vermessungs- und Geoinformationwesens*. Berlin, New York: Walter de Gruyter. 34, 39, 40, 45
- Oggier, T., Büttgen, B. & Lustenberger, F. (2005). *SwissRanger SR3000 and first experiences based on miniaturized 3D-TOF Cameras* (Bericht). Swiss Center for Electronics and Microtechnology (CSEM). 108
- Oggier, T., Lehmann, M., Kaufmann, R., Schweizer, M., Richter, M., Metzler, P. et al. (2004, feb). An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). In L. Mazuray, P. J. Rogers & R. Wartmann (Hrsg.), *Optical Design and Engineering* (Bd. 5249, S. 534-545). 110
- Ohno, K., Nomura, T. & Tadokoro, S. (2006). Real-time robot trajectory estimation and 3D map construction using 3D camera. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2006* (S. 5279-5285). 112
- Oikonomidis, I., Kyriazis, N. & Argyros, A. (2011). Efficient model-based 3D tracking of hand articulations using kinect. In *Proceedings of the 22nd British Machine Vision Conference, BMVC 2011*. Dundee, UK: BMVA. 243, 244
- OpenKinect. (2011). *OpenKinect Project*. Zugriff am 05.10.2011 auf <http://openkinect.org/> 243
- Oprisescu, S., Falie, D., Ciuc, M. & Buzuloiu, V. (2007). Measurements with ToF cameras and their necessary corrections. In *Proceedings of ISSCS 2007*. 123
- Oswald, W. (1992). *Zur kombinierten Ausgleichung heterogener Beobachtungen in hybriden Netzen* (Schriftenreihe Studiengang Vermessungswesen Nr. 44). Universität der Bundeswehr München. 45
- Parlitz, C., Baum, W., Reiser, U. & Hägele, M. (2007). Intuitive human-machine-interaction and

- implementation on a household robot companion. In *Proceedings of the 2007 Conference on Human interface* (S. 922–929). Berlin, Heidelberg: Springer-Verlag. Verfügbar unter <http://portal.acm.org/citation.cfm?id=1766451.1766560> 113
- Patterson, M. L. (1983). *Nonverbal Behavior – A Functional Perspective*. Springer-Verlag. 225
- Pattinson, T. (2010a). *Quantification and Description of Distance Measurement Errors of a Time-of-Flight Camera*. Master thesis, University of Stuttgart, Stuttgart, D. 117, 118
- Pattinson, T. (2010b). Quantification and description of distance measurement errors of a time-of-flight camera. In *DAGM Young Researchers Forum 2010*. Darmstadt. 122, 163
- Penven, A. (2011). *Dancing With Invisible Light*. Zugriff am 04.10.2011 auf <http://audreypenven.net/> 242
- Piironen, T. H., Paul, D. & Kontio, P. (2000). *Arrangement and Method for Inspection of Surface Quality* (Nr. EP20000660024). 70
- Pikus, I. (2010). *3DV Systems ZCam Follows Gamer Movements and Images 3D Style*. Zugriff am 14.09.2010 auf <http://www.besportier.com/archives/3dv-systems-zcam-follows-gamer-1.html> 112
- PMDTec. (2010). Datasheet PMD [vision] CamCube 3.0 [Software-Handbuch]. Siegen. Verfügbar unter <http://www.pmdtec.com/> 24, 103, 104, 106, 118, 125, 144
- Pradhan, R., Ghose, M. K. & Jeyaram, A. (2010, October). Extraction of depth elevation model (DEM) from high resolution satellite imagery using shape from shading approach. *International Journal of Computer Applications*, 7 (12), 40–46. 70
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1992). *Numerical Recipes in C: The Art of Scientific Computing* (2. Aufl.). Cambridge University Press. 153
- Purde, A. (2006). *Speckle-Interferometrie zur Formvermessung unstetiger Oberflächen*. Dissertation, Technische Universität München, München. 66
- Putze, T. (2004). Einsatz einer Highspeedkamera zur Bestimmung von Geschwindigkeitsfeldern in Gasströmungen. In *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation*. 89
- Putze, T. (2008). Photogrammetrische Erfassung von orthopädischen Laufparametern zur Ganganalyse. In T. Luhmann (Hrsg.), *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik (Beiträge Oldenburger 3D-Tage 2008)*. Verlag Herbert Wichmann. 78
- Radmer, J., Fuste, P. M., Schmidt, H. & Kruger, J. (2008). Incident light related distance error study and calibration of the pmd-range imaging camera. *Computer Vision and Pattern Recognition Workshop*, 1-6. 124
- Ramseyer, F. (2010). Nonverbale Synchronisation in therapeutischer Interaktion. In S. Bender (Hrsg.), *Bewegungsanalyse von Interaktionen* (S. 11-32). Berlin, D: Logos. 20, 22
- Rapp, H. (2007). *Experimental and theoretical investigation of correlating TOF-camera systems*. Diplomarbeit, Universität Heidelberg, Insitut für Wissenschaftliches Rechnen, Heidelberg, D. 117
- Reißmann, G. (1976). *Die Ausgleichsrechnung - Grundlagen und Anwendungen in der*

- Geodäsie*. Berlin: VEB Verlag für Bauwesen. 34, 43
- Reiser, U. & Kubacki, J. (2007, June). Using a 3D time-of-flight range camera for visual tracking. In *Symposium on Intelligent Autonomous Vehicles (IAV)*. Toulouse, F. 112, 188
- Remondino, F. (2003). Recovering metric information from old monocular video sequences. In G. A. & K. H. (Hrsg.), *6th Conference on Optical 3D Measurement Techniques* (S. 214-222). Zurich, CH: ETH, Swiss Federal Institute of Technology, Institute of Geodesy and Photogrammetry. 178
- Remondino, F. (2004, July 7-9). Human motion reconstruction and animation from video sequences. In *17th International Conference on Computer Animation and Social Agents (CASA2004)* (S. 347-354). Geneva, CH: Computer Graphics Society (CGS). 178
- Remondino, F. (2006). *Image-based Modeling for Object and Human Reconstruction*. Ph.d. thesis, nr. 16562, Technische Wissenschaften, Eidgenössische Technische Hochschule ETH Zürich. 176, 179
- Reulke, R. (2006). Combination of distance data with high resolution images. In *ISPRS Commission V Symposium*. 128
- Riegl. (2010). Extra long range & high accuracy 3d terrestrial laser scanner LSM-Z620 [Software-Handbuch]. Verfügbar unter <http://www.riegl.com> 75
- Ringbeck, T. & Hagebecker, B. (2007). Dreidimensionale Objekterfassung in Echtzeit - PMD Kameras erfassen pro Pixel Distanz und Helligkeit mit Videoframerate. *Allgemeine Vermessungs-Nachrichten*, 114 (7), 263-270. 112, 119
- Ringbeck, T., Möller, T. & Hagebecker, B. (2007). Multidimensional measurement by using 3-D PMD sensors. *Advances in Radio Science*, 5, 135-146. 106, 110, 118
- Robbins, S., Murawski, B. & Schröder, B. (2009). Photogrammetric calibration and colorization of the swissranger SR-3100 3-D range imaging sensor. In *Optical engineering* (Bd. 48, S. 053603-053603-8). 129
- Robbins, S., Schröder, B., Murawski, B., Heckman, N. & Leung, J. (2008). Photogrammetric calibration of the SwissRanger 3D range imaging sensor. In *Proc. SPIE: Optical Sensors 2008* (Bd. 7003, S. 7003-20). 129
- Rodehorst, V. & Hellwich, O. (2006). Genetic algorithm sample consensus (GASAC) - a parallel strategy for robust parameter estimation. In *CVPRW '06: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop* (S. 103). Washington, DC, USA: IEEE Computer Society. 51
- Rost, C. (2011). *Phasenmehrwegereduzierung basierend auf Signalqualitätsmessungen geodätischer GNSS-Empfänger*. Dissertation, Technische Universität Dresden, Geodätisches Institut, Dresden, D. 119
- Santrac, N., Friedland, G. & Rojas, R. (2006). *High resolution segmentation with a time-of-flight 3D-camera using the example of a lecture scene* (Technical Report Nr. B-06-09). Freie Universität Berlin, Fachbereich Mathematik und Informatik. 128
- Scheller, S. & Schneider, D. (2006). Extraktion von Primitiven aus Laserscannerpunktwolken zur

- Rekonstruktion von Tragwerken. In T. Luhmann (Hrsg.), *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik (Beiträge Oldenburger 3D-Tage 2006)*. Heidelberg: Herbert Wichmann Verlag. 75
- Schiller, I., Beder, C. & Koch, R. (2008). Calibration of a PMD-camera using a planar calibration pattern together with a multi-camera setup. In J. J. Chen J. & H.-G. Maas (Hrsg.), *International archives of the photogrammetry, remote sensing and spatial information sciences* (Bd. XXXVII, S. 297ff). 122, 129, 132
- Schmidt, M., Jehle, M. & Jähne, B. (2008). Range flow estimation based on photonic mixing device data. *IJISTA*, 5 (3/4), 380-392. 191, 193
- Schmidt, M. O. (2008). *Spatiotemporal Analysis of Range Imagery*. Dissertation, Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg, Heidelberg, Germany. 125, 191, 192, 194
- Schmitter, D. (2005). *Ermitteln der körperlichen Belastung bei Tätigkeiten im Sitzen* (Mitteilung Nr. 88212.d). Luzern, CH: Schweizerische Unfallversicherungsanstalt (SUVA). 176, 177
- Schneider, D. (2009). *Geometrische und stochastische Modelle für die integrierte Auswertung terrestrischer Laserscannerdaten und photogrammetrischer Bilddaten*. Dissertation, Technische Universität Dresden, Fakultät Forst-, Geo- und Hydrowissenschaften, Professur für Photogrammetrie. 45, 58, 60, 63, 75, 131
- Schulze, M. (2008). *Implementation und Validierung von Templateupdate-Strategien in der Bildsequenzanalyse*. Unveröffentlichte Studienarbeit, Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Dresden. 89
- Schulze, M. (2010). 3D-camera based navigation of a mobile robot in an agricultural environment. In J. P. Mills, D. M. Barber, P. Miller & I. Newton (Hrsg.), *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII, S. 538-542). 49, 112
- Schulze, M., Nöthe, M., Kieback, B. & Maas, H.-G. (2005). Material analysis applying microfocus computer tomography and 3D photogrammetry. In *Deutsche Physikalische Gesellschaft, Frühjahrstagung*. Berlin, D. 98
- Schwarte, R. (1996). Eine neuartige 3D-Kamera auf der Basis eines 2D-Gegentaktkorrelator-Arrays. In *Aktuelle Entwicklungen und industrieller Einsatz der Bildverarbeitung* (S. 111-117). Aachen: MIT GmbH. 105
- Schwarte, R. (1998). *Verfahren und Vorrichtung zur Bestimmung der Phasen- und/oder Amplitudeninformation einer elektromagnetischen Welle* (Nr. DE 197 04 496). 105
- Schwarte, R., Heinol, H., Buxbaum, B., Ringbeck, T., Xu, Z. & Hartmann, K. (1999). Principles of three-dimensional imaging techniques. In B. Jähne, H. Haußecker & P. Geißler (Hrsg.), *Handbook of Computer Vision and Applications - Sensors and Imaging* (Bd. 1, S. 463-484). Academic Press. 65, 66, 103
- Schwarz, L., Mateus, D., Castaneda, V. & Navab, N. (2010). Manifold learning for ToF-based human body tracking and activity recognition. In *Proceedings of the British Machine Vision*



- Conference* (S. 80.1-80.11). BMVA Press. 243
- Shapiro, L. G. & Haralick, R. M. (1987). Relational matching. *Applied Optics*, 26, 1845-1851. 86
- Sheh, R., Kadous, M. W. & Sammut, C. (2006). *On Building 3D Maps Using a Range Camera: Applications to Rescue Robotics* (Technical Report Nr. UNSW-CSE-TR-0609). Sydney, Australia: School of Computer Science and Engineering, The University of New South Wales. 112
- Shi, J. & Tomasi, C. (1994). Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition* (S. 593-600). 84
- Shotton, J. & Sharp, T. (2011). Real-time human pose recognition in parts from single depth images. In *IEEE Conference on Computer Vision and Pattern Recognition*. 243, 244
- Sidenbladh, H., Black, M. J. & Fleet, D. J. (2000). Stochastic tracking of 3D human figures using 2D image motion. In *Proceedings of the 6th European Conference on Computer Vision* (S. 702-718). London, UK: Springer-Verlag. 178
- Sieg, D. & Hirsch, M. (2000a). Varianzkomponentenschätzung in ingenieurgeodätischen Netzen, Teil 1: Theorie. *Allgemeine Vermessungsnachrichten*, 3, 82-90. 45, 47
- Sieg, D. & Hirsch, M. (2000b). Varianzkomponentenschätzung in ingenieurgeodätischen Netzen, Teil 2: Anwendungen. *Allgemeine Vermessungsnachrichten*, 4, 122-137. 45
- SIMI Reality Motion Systems GmbH. (2011). *2D/3D Bewegungsanalyse und Verhaltensanalyse*. Unterschleißheim, D. Zugriff am 23.03.2011 auf <http://www.simi.com/de/> 182, 183
- Smith, M. L. & Smith, L. N. (2005). Dynamic photometric stereo - a new technique for moving surface analysis. *Image Vision Comput.*, 23, 841-852. 70
- Smith, W. A. & Hancock, E. R. (2010). Estimating facial reflectance properties using shape-from-shading. *Int. J. Comput. Vision*, 86, 152-170. 70
- Snow, K. B. (2002). *Applications of Parameter Estimation and Hypothesis Testing of GPS Network Adjustments* (Bericht Nr. 465). Geodetic and GeoInformation Science, Department of Civil Environmental Engineering and Geodetic Science, The Ohio State University, Columbus, Ohio, USA. 39, 155
- Soules, L. & Wiens, K. (2011). *Microsoft Kinect Teardown*. Zugriff am 04.10.2011 auf [www.ifixit.com/Teardown/Microsoft-Kinect-Teardown/4066/](http://www.ifixit.com/Teardown/Microsoft-Kinect-Teardown/4066/) 242
- Spies, H., Jähne, B. & Barron, J. L. (2002). Range flow estimation. *Computer Vision and Image Understanding*, 85 (3), 209-231. 191
- Spirig, T., Seitz, P., Vietze, O. & Heitger, F. (1995). The lock-in CCD-two-dimensional synchronous detection of light. *IEEE Journal of Quantum Electronics*, 31 (9), 1705-1708. 105
- Stoffert, G. (1985). Analyse und Einstufung von Körperhaltungen bei der Arbeit nach der OWAS-Methode. *Zeitschrift für Arbeitswissenschaft*, 39 (11NF) (1), 31-38. 175
- sueddeutsche.de. (2011). *Softwarehersteller Microsoft - Rekordumsatz - der Konsole sei Dank*. Zugriff am 04.10.2011 auf <http://www.sueddeutsche.de/wirtschaft/softwarehersteller-microsoft-rekordumsatz-der-konsole-sei-dank-1.1052287> 241

- Swadzba, A., Beuter, N., Wachsmuth, S. & Kummert, F. (2010). Dynamic 3D scene analysis for acquiring articulated scene models. In *International conference on robotics and automation*. Anchorage, AK, USA: IEEE. 189, 190
- Thierfelder, S., Gossow, D., Navarro Luzón, C., Nowack, S., Merten, N., Friedmann, S. et al. (2011). *RoboCup 2011 - homer@UniKoblenz (Germany)* (Bericht). Universität Koblenz-Landau. 243, 244
- Tomasi, C. & Kanade, T. (1991). *Detection and Tracking of Point Features* (Bericht). International Journal of Computer Vision. 84
- Torr, P. & Zisserman, A. (2000). Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78 (1), 138-156. 51
- Urtasun, R. & Fua, P. (2004). 3-D human body tracking using deterministic temporal motion models. In *European Conference on Computer Vision, Prague, Czech Republic*. 178
- Vedula, S., Baker, S., Rander, P., Collins, R. & Kanade, T. (2005, March). Three-dimensional scene flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 (1), 475 - 480. 180, 181
- Vicon House of Moves, Inc. (2008). *VICON House of Moves Builds New Motion Capture Sound Stage; Expands Staff with Full Service Animation Team*. Los Angeles, USA. Zugriff am 23.03.2011 auf <http://blog.digitalcontentproducer.com/briefingroom/2008/07/29/> 183
- Vock, D., Gumhold, S., Spehr, M., Westfeld, P. & Maas, H.-G. (2010). GPU-based volumetric reconstruction of trees from multiple images. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII, S. 586-591). Newcastle upon Tyne, UK. 71, 72
- Vosselman, G. (1992). *Relational Matching* (Bd. 628). Springer. 86
- Vosselman, G. & Maas, H.-G. (Hrsg.). (2010). *Airborne and Terrestrial Laser Scanning*. Caithness, UK: Whittles Publishing. 75
- Wanninger, L. & May, M. (2000). Carrier phase multipath calibration of GPS reference stations. In *Proc. of ION GPS 2000* (S. 132-144). Salt Lake City, UT. 119
- Webb, E. J., Campbell, D. T., Schwartz, R. D. & Sechrest, L. (1975). *Nicht-reaktive Meßverfahren*. Weinheim, D: Beltz. 20, 175
- Weingarten, J. W., Gruener, G. & Siegwari, R. (2004). A state-of-the-art 3D sensor for robot navigation. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems* (S. 2155–2160). 112
- Westfeld, P. (2005). *Entwicklung von Verfahren zur räumlich und zeitlich aufgelösten 2D-Bewegungsanalyse in der Glaziologie*. Unveröffentlichte Diplomarbeit, Technische Universität Dresden, Institut für Photogrammetrie und Fernerkundung, Dresden, D. 88, 91, 96, 98, 100
- Westfeld, P. (2007a). Ansätze zur Kalibrierung des Range-Imaging-Sensors SR-3000 unter simultaner Verwendung von Intensitäts- und Entfernungsbildern. In T. Luhmann (Hrsg.), *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik (Beiträge Oldenburger 3D-Tage 2007)* (S. 137-146). Herbert Wichmann Verlag Heidelberg. 128, 130, 133



- Westfeld, P. (2007b). Development of approaches for 3-D human motion behaviour analysis based on range imaging data. In A. Grün & H. Kahmen (Hrsg.), *Optical 3-D Measurement Techniques VIII* (Bd. II, S. 393-402). Zurich. 113, 192, 193, 226, 227, 228, 229
- Westfeld, P. & Hempel, R. (2008). Range image sequence analysis by 2.5-D least squares tracking with variance component estimation and robust variance covariance matrix estimation. In J. J. Chen J. & H.-G. Maas (Hrsg.), *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVII, S. 457-462). 191, 192, 211
- Westfeld, P. & Maas, H. G. (2006). *Unveröffentlichte Studie zur Einsetzbarkeit von Range Cameras bei der räumlichen Abbildung von Abwasserbauwerken*. 113
- Westfeld, P. & Maas, H.-G. (2010). 3-D least squares tracking in time-resolved tomographic reconstruction of dense flow marker fields. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Bd. XXXVIII, S. 597-602). Newcastle upon Tyne, UK. 100
- Westfeld, P., Mulsow, C. & Schulze, M. (2009). Photogrammetric calibration of range imaging sensors using intensity and range information simultaneously. *Optical 3D Measurement Techniques IX, Vol. II*, pp. 129. 49, 130, 133
- Willneff, J. & Grün, A. (2002). A new spatio-temporal matching algorithm for 3D-particle tracking velocimetry. In *9th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*. Honolulu, Hawaii. 180
- Winkelbach, S. (2006). *Das 3D-Puzzle-Problem - Effiziente Methoden zum paarweisen Zusammensetzen von dreidimensionalen Fragmenten*. Dissertation, Technischen Universität Carolo-Wilhelmina zu Braunschweig, Carl-Friedrich-Gauß-Fakultät für Mathematik und Informatik, Braunschweig. 68
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)* (S. 1556-1559). 177
- Woodham, R. (1978). Photometric stereo: A reflectance map technique for determining surface orientation from image intensity. In *SPIE Proceedings Series* (Bd. 155, S. 136-143). 70
- Wujanz, D., Weisbrich, S. & Neitzel, F. (2011). 3D-Mapping mit dem Microsoft Kinect Sensor - erste Untersuchungsergebnisse. In T. Luhmann & C. Müller (Hrsg.), *Photogrammetrie - Laserscanning - Optische 3D-Messtechnik (Beiträge Oldenburger 3D-Tage 2011)* (S. 274-283). Verlag Herbert Wichmann. 243, 244
- Xu, Z., Schwarte, R., Heinol, H., Buxbaum, B. & Ringbeck, T. (2005). *Smart pixel - photonic mixer device (PMD) new system concept of a 3D-imaging camera-on-a-chip* (Bericht). PMDTechnologies GmbH. 105
- Yahav, G., Iddan, G. J. & Mandelbom, D. (2007). Camera for gaming application using a novel 3D imager. In A. Grün & H. Kahmen (Hrsg.), *Optical 3-D Measurement Techniques VIII* (Bd. I, S. 42-47). Zurich. 103, 104, 113
- Yamamoto, M., Boulanger, P., Beraldin, J. A. & Rioux, M. (1993). Direct estimation of range flow

- on deformable shape from a video rate range camera. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15, 82–89. 191
- York-Barr, J. & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research*, 74 (3), 255-316. 175
- Zalevsky, Z., Shpunt, A., Maizels, A. & Garcia, J. (2007). *Method and System for Object Reconstruction* (Nr. WO 2007/043036 A1). Hertzelia, IL. 242
- Zhang, Y. & Kambhamettu, C. (2000). Integrated 3D scene flow and structure recovery from multiview image sequences. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Bd. 2, S. 2674-2681). 191
- Zywitzka, F., Massen, J., Brunn, M., Lang, C. & Görnig, T. (2005). One-to-three-dimensional ranging for future automotive safety systems. In H. Ingensand & T. Kahlmann (Hrsg.), *Range Imaging Research Day* (Bd. 1). Zurich. (ISBN 3-906467-57-0) 112



# Abbildungsverzeichnis

- 1.1 Zwischenmenschliches Verhalten – 19
- 1.2 Manuelle Erfassung eines Interaktionsverlaufs – 21
- 1.3 Interaktive Erfassung eines Interaktionsverlaufs – 21
- 1.4 Automatisierte Erhebung durch Trackingsysteme – 22
- 1.5 Automatisierte Erhebung durch MEA – 22
- 1.6 3D-Kamerabildsequenzanalyse in der Verhaltensforschung – 24
  
- 4.1 Ebene Ähnlichkeitstransformation – 32
- 4.2 Räumliche Ähnlichkeitstransformation – 33
- 4.3 Kurvenverlauf und Parameter einer Normalverteilung von Zufallsgrößen – 35
- 4.4 Prinzip des RANSAC-Algorithmus am Beispiel einer Geradeneinpassung – 50
  
- 5.1 Photogrammetrisches Messprinzip – 54
- 5.2 Referenzfeldkalibrierung – 60
- 5.3 Aufnahmekonfiguration zur Selbstkalibrierung – 62
  
- 6.1 Klassifikation berührungsloser 3D-Messverfahren – 65
- 6.2 Triangulationsverfahren – 66
- 6.3 Prinzip der aktiven optischen Triangulation – 67
- 6.4 Prinzipien des Lichtschnittverfahrens und der strukturierten Beleuchtung – 68
- 6.5 Anwendungen der photogrammetrischen Bildzuordnung – 69
- 6.6 Schattierungsverfahren – 71
- 6.7 Silhouettenschnittverfahren – 72
- 6.8 Verfahren zur Laufzeitmessung – 74
- 6.9 Messprinzip des Impulsmessverfahrens – 74
- 6.10 Messprinzip des Phasenvergleichsverfahrens – 74
- 6.11 3D-Laserscanning – 75
  
- 7.1 Anwendungen im Bereich der automatischen Bewegungsanalyse – 78
- 7.2 Reale vs. visuell wahrnehmbare Objektbewegung – 79
- 7.3 3D-Bildblock – 80
- 7.4 Differenzbildanalyse – 81
- 7.5 Blendenproblem – 87

- 7.6 Korrespondenzproblem – 87
- 7.7 Prinzip der Kreuzkorrelation – 90
- 7.8 Optisches Flussfeld eines Zylinders und einer Kugel – 92
- 7.9 Parameter einer 2D-Affintransformation und ihre Auswirkungen im Bild – 99
- 7.10 3D-Affintransformation eines Kuboids – 99
- 7.11 Oberflächenausschnitte zweier Laserscannerstreifen in einer TIN-Struktur – 99
  
- 8.1 Daten einer 3D-Kamera – 104
- 8.2 Aktuelle 3D-Kameramodelle – 104
- 8.3 Der PMD-Halbleiterchips als Kernstück einer 3D-Kamera – 105
- 8.4 Systemkomponenten eines SwissRanger™ SR3k – 108
- 8.5 Analyse der AKF mittels Phasenvergleich – 108
- 8.6 Abbildungsgeometrie einer 3D-Kamera – 111
- 8.7 Anwendungsbeispiele einer 3D-Kamera, I – 112
- 8.8 Anwendungsbeispiele einer 3D-Kamera, II – 113
  
- 9.1 Einlaufverhalten einer 3D-Kamera – 118
- 9.2 Auswirkungen des *Scattering*-Effekts – 120
- 9.3 Soll-Ist-Abweichungen – 122
- 9.4 Einfluss der Amplitude auf die Entfernungsmessung – 124
- 9.5 Visualisierung von Bewegungsartefakten – 125
  
- 10.1 Klassifikation von 3D-Kamerakalibrieransätzen – 127
- 10.2 Photogrammetrische 3D-Kamerakalibrierung – 132
- 10.3 Distanzkalibrierung einer 3D-Kamera – 132
- 10.4 Simultane Kalibrierung einer 3D-Kamera – 132
  
- 11.1 Geometrisches Modell der integrierten Bündelblockausgleichung – 134
- 11.2 Datenfluss für die integrierte Bündelblockausgleichung – 136
- 11.3 Amplitudenbildpunktmessung mit 2D-LSM – 138
- 11.4 Translationsparametergenauigkeit – 139
- 11.5 Schematische Darstellung der sequentiellen Kugelsegmentierung – 140
- 11.6 Kugelsegmentierung im Amplituden- und Entfernungsbild – 141
- 11.7 Grobe Lokalisierung – 142
- 11.8 Histogrammanalyse im Amplitudenbild – 142
- 11.9 Sternförmige Profilanalyse im Entfernungsbild – 143
- 11.10 Programmstruktur der integrierten Bündelblockausgleichung – 156
- 11.11 Referenzfeld – 159
- 11.12 Beispiel für die Netzkonfiguration eines 3D-Kamerabildverbands – 161
- 11.13 Referenzfeldaufnahmen – 161

- 11.14 Parameter der äußeren Orientierung – 162
- 11.15 Auswirkungen der Distanzkorrektur  $\Delta D$  – 164
- 11.16 Residuen der Bildkoordinatenmessungen – 166
- 11.17 Residuen der Distanzmessungen – 167
- 11.18 Äußere Genauigkeit – 168
  
- 12.1 Interaktive Erfassung und manuelle Interpretation von Bewegungsabläufen – 177
- 12.2 2D-Punktverfolgung in Einzelbildsequenzen – 179
- 12.3 Körperrekonstruktion aus Einzelbildern – 179
- 12.4 Spatiotemporale Silhouettenbildzuordnung – 179
- 12.5 Menschliche Körper- und Bewegungsrekonstruktion – 181
- 12.6 3D-Bewegungsanalyse einer gehenden Person – 181
- 12.7 3D-Bewegungsanalyse einer tanzenden Person – 181
- 12.8 Softwarepakete zur Punktverfolgung – 183
- 12.9 *Motion Capturing Studio* – 183
- 12.10 Bewegungserfassung in Filmindustrie und Sportwissenschaft – 183
  
- 13.1 Personendetektion mit Partikelfilter – 187
- 13.2 3D-Kamera- und Silhouetteninformation – 187
- 13.3 Visuelle Personenverfolgung mit Partikelfilter – 187
- 13.4 Ergebnisse einer modellbasierten Gesichtsrekonstruktion – 190
- 13.5 Körpermodellierung – 190
- 13.6 Bewegungsanalysen in 3D-Kamerabildsequenzen zur Umgebungskartierung – 190
- 13.7 Gestendetektion und -interpretation auf der Basis von 3D-Kameradaten – 190
- 13.8 Ganganalysen am Laufband auf der Basis von 3D-Kameradaten – 194
- 13.9 *Range Flow* einer Kugel – 194
- 13.10 Visualisierung der Eingabedaten und der Ergebnisse eines 2,5D-LST – 194
  
- 14.1 Zusätzliche Neigungsparameter und ihre Auswirkungen im Bild – 197
- 14.2 Bildmaßstabszahl – 200
- 14.3 Maximal abbildbare Neigung um die  $x'$ -Achse – 202
- 14.4 Programmstruktur des integrierten 2,5D-LST – 207
- 14.5 Simulierte Daten – 213
- 14.6 Statisches Referenzobjekt – 213
- 14.7 Winkbewegung – 213
- 14.8 Verfolgung einer Handbewegung – 214
- 14.9 Residuenverteilung, I – 219
- 14.10 Residuendarstellung, I – 219
- 14.11 Residuenverteilung, II – 220
- 14.12 Residuendarstellung, II – 220

- 
- 14.13 Plot des Konvergenzverhaltens – 221
  - 14.14 Iterative Anpassung der Suchbildausschnitte – 221
  
  - 15.1 Konflikte unter Kindern in Kindertagesstätten – 226
  - 15.2 Berechnung der interpersonalen Distanz aus 3D-Kameradaten – 227
  - 15.3 Statistische Modellierung der interpersonalen Distanz aus RIM-Daten – 227
  - 15.4 Berechnung der Körperorientierung aus 3D-Kameradaten – 229
  - 15.5 Automatische Verfolgung der Schulterpunkte durch eine 3D-Kamerabildsequenz – 229
  - 15.6 Gestenerkennung – 230
  - 15.7 Automatische Detektion, Markierung, Klassifikation sprachbegleitender Gesten – 232
  - 15.8 3D-Identifizierung des Bewegungsorts einer Geste – 233
  
  - 17.1 Kinect-Sensorleiste – 242
  - 17.2 Funktionsprinzip des Kinect-Sensors – 242
  - 17.3 NIR-Specklemuster – 242
  - 17.4 Bewegungssteuerung von Konsolenspielen – 244
  - 17.5 Daten des Kinect-Sensors – 244
  - 17.6 Anwendungsbeispiele – 244



## Tabellenverzeichnis

- 8.1 Vergleich von Verfahren zur dynamischen 3D-Oberflächenerfassung – 115
- 10.1 Photogrammetrischen 3D-Kamerakalibrierung – 128
- 11.1 Vereinbarungen zur Indizierung von Beobachtungen und Unbekannten – 135
- 11.2 Synthetisches Beispiel zur Ermittlung der Anzahl der Beobachtungen – 143
- 11.3 Synthetisches Beispiel zur Ermittlung der max. Anzahl unbekannter Parameter – 147
- 11.4 Synthetisches Beispiel zur Ermittlung der maximalen Anzahl an Bedingungen – 150
- 11.5 Parameter der inneren Orientierung – 162
- 11.6 Parameter der Distanzkorrekturterme und ihre a-posteriori Standardabweichungen – 164
- 11.7 Beobachtungsfehler – 166
- 11.8 Vergleich des sequentiellen mit dem simultanen Ansatz – 170
- 14.1 2,5D-LST-Genauigkeitsangaben – 215
- 14.2 Prozentsatz signifikant bestimmter Transformationsparameter – 217



## Abkürzungsverzeichnis

2,5D	Zweieinhalbdimensional
2D	Zweidimensional
3D	Dreidimensional
AKF	Autokorrelationsfunktion
ALS	<i>Airborne</i> Laserscanner
API	Schnittstelle zur Anwendungsprogrammierung
APS	Aktiver Pixelsensor
BCCE	<i>Brightness Constancy Constraint Equation</i>
CCD	<i>Charge-Coupled Device</i>
CMOS	<i>Complementary Metal Oxide Semiconductor</i>
CONDENSATION	<i>Conditional Density Propagation for Visual Tracking</i>
DFT	<i>Diskrete Fouriertransformation</i>
DLT	Direkte lineare Transformation
FACS	<i>Facial Action Coding System</i>
FFT	<i>Fast Fourier Transformation</i>
FPN	<i>Fixed Pattern Noise</i>
GASAC	<i>Genetic Algorithm Sampling Consensus</i>
GHM	Gauß-Helmert-Modell
GMM	Gauß-Markov-Modell
GPS	Globales Positionierungssystem
GWDS	Grenzwert für das <i>Data Snooping</i>
HCI	<i>Human-Computer Interaction</i>
HKA	Hauptkomponentenanalyse
ICP	<i>Iterative Closest Point Algorithm</i>

---

IO	<i>Interior Orientation</i>
LED	Lichtemittierende Diode
LIDAR	<i>Light Detection and Ranging</i>
LO-RANSAC	<i>Local Optimised Random Sample Consensus</i>
LS3D	<i>Least Squares 3D Surface Matching</i>
LSM	<i>Least Squares Matching</i>
LST	<i>Least Squares Tracking</i>
LUT	Lookup-Tabelle
MEA	<i>Motion Energy Analysis</i>
MKQ	Methode der kleinsten Quadrate
MSAC	<i>M-Estimator Sample Consensus</i>
NEUROGES	<i>Neuropsychological Gesture Coding System</i>
NIR	Nahes Infrarot
PMD	Photomischdetektor
PSF	Punktspreizfunktion
RANSAC	<i>Random Sample Consensus</i>
RFCE	<i>Range Flow Constraint Equation</i>
RGB	Rot, Grün, Blau
RGBD	Rot, Grün, Blau, Distanz
RIM	<i>Range Imaging</i>
RMS	<i>Root Mean Square Error</i>
SBI	<i>Suppression of Background Illumination</i>
SDK	<i>Software Development Kit</i>
SfS	<i>Shape from Shading</i>
SMB	<i>Suppression of Motion Blur</i>
SNR	Signal-Rausch-Verhältnis
SPAD	<i>Single Photon Avalanche Dioden</i>
TIN	<i>Triangulated Irregular Network</i>
TLS	Terrestrischer Laserscanner
ToF	<i>Time-of-Flight</i>
USB	<i>Universal Serial Bus</i>

---

VGA	<i>Video Graphics Array</i>
VHS	<i>Video Home System</i>
VKS	Varianzkomponentenschätzung



## Danke

Lutz-Michael Alisch Uwe Altmann Michael Bauer Cornelia Becker Ditte Becker Heike Becker Anne Bienert Davide Bonmassar Oliver Bringmann Sebastian Drichelt Oliver Fucke Dorian Gärtner Renate Härtel Juliane Hanka Maria Hempel René Hempel Rico Hermkes Nalani Hetti Nicole Hey Marcus Hillwig Max Hofmann Carola Hohnvehlmann Marleen Hohnvehlmann Michael Hohnvehlmann Alexander Hoppe Claudia Hoppe Peter Joraschky Jörg Knorr Tobias Kohoutek Nadine König Robert Koschitzki Danny Krebs Hedda Lausberg Stefan Lischke Hans-Gerd Maas Tolja Meissner Angelina Morgenstern Christian Mulsow Kai Nehring Elisabeth Ostwald Katharina Pech Sanka Perera Torsten Putze Sigrid Pönitz Ralf Reulke Anette Richter Katja Richter René Rohde Christian Rost Paul Karl Rost Nadine Rülke Mehran Sattari Steffen Scheller Anita Schilling Anja Schmidt Danilo Schneider Nancy Schneider Marc Schulze Matthias Schulze Ellen Schwalbe Christian Schönau Linda Schönau Luca Schönau Nadine Stelling Susana García Rodríguez Cristina Tapia Rodríguez Sylvia Trezdziaik Edith Vorwerk Rudolf Vorwerk Peter Welchman Günter Westfeld Irmgard Westfeld Werner Westfeld