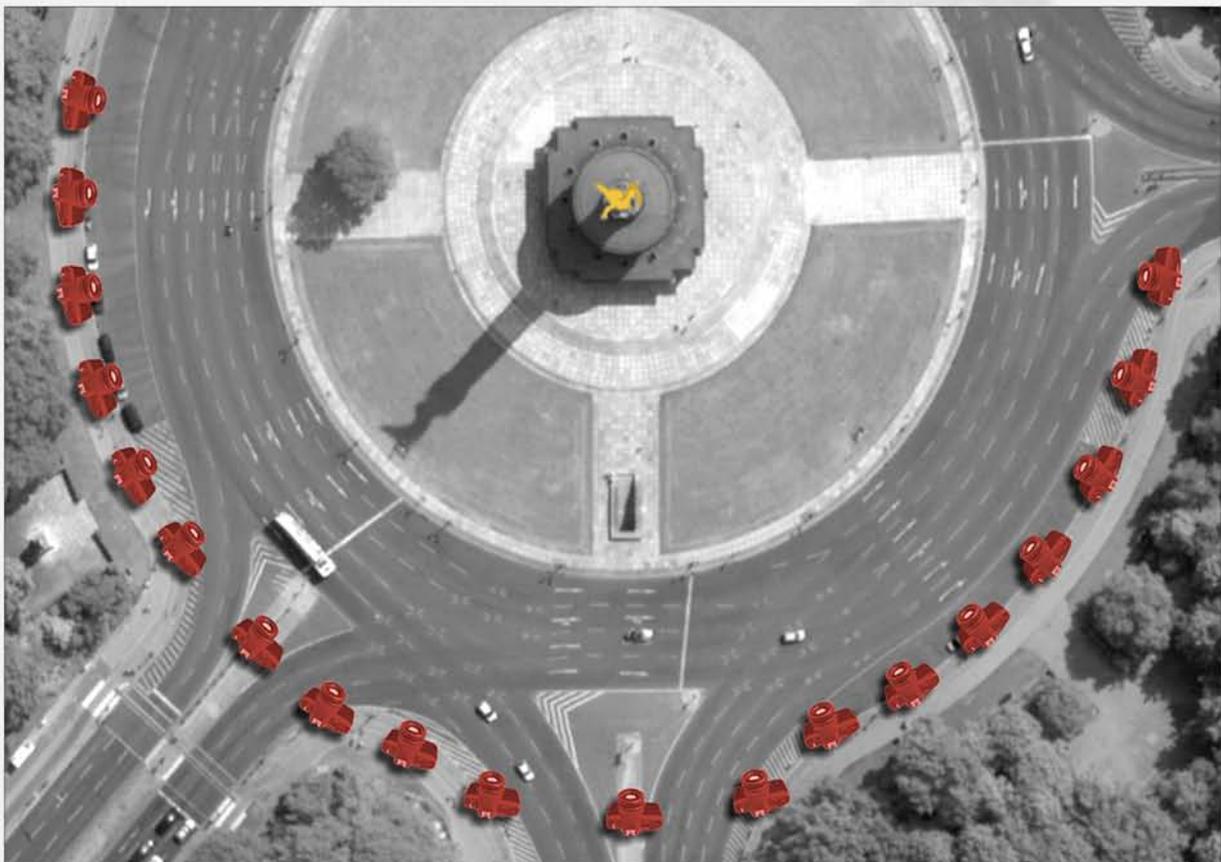


Volker Rodehorst



# Photogrammetrische 3D-Rekonstruktion

im Nahbereich durch Auto-Kalibrierung  
mit projektiver Geometrie



Wissenschaftlicher Verlag Berlin  
**wvb**



# **Photogrammetrische 3D-Rekonstruktion im Nahbereich durch Auto-Kalibrierung mit projektiver Geometrie**

VOLKER RODEHORST

von der Fakultät VI  
Bauingenieurwesen und Angewandte Geowissenschaften  
der Technischen Universität Berlin

zur Erlangung des akademischen Grades eines  
Doktors der Ingenieurwissenschaften (Dr.-Ing.)  
genehmigte Dissertation

Gutachter: Prof. Dr.-Ing. OLAF HELLWICH  
Gutachter: Prof. Dr.-Ing. JÖRG ALBERTZ  
Gutachter: Prof. Dr. sc. nat. REINHARD KLETTE  
Vorsitzender: Prof. Dr.-Ing. LOTHAR GRÜNDIG

Tag der wissenschaftlichen Aussprache: 19. Dezember 2003

Berlin 2004

D 83



## Kurzfassung

Für die Rekonstruktion von dreidimensionalen Objekten aus photographischen Bildern wird die Position und Orientierung der verwendeten Aufnahmesysteme benötigt. Da aber häufig nur das Bildmaterial zur Verfügung steht, muss die zugrunde liegende Abbildungsgeometrie mit Hilfe einer Selbst-Kalibrierung bestimmt werden. Dabei stellen Anwendungen im Nahbereich aufgrund der flexiblen Bildanordnungen und der perspektivischen Verzerrungen besonders hohe Anforderungen an ein automatisches Auswerteverfahren.

In dieser Arbeit wird durch die Integration und Verbesserung von aktuellen Forschungsergebnissen ein durchgängiger Auswerteprozess vorgeschlagen. Die Schwerpunkte liegen in der Entwicklung von Beiträgen zu einer merkmalsbasierten Korrespondenz-Analyse und einer optimierten Auto-Kalibrierung unter Verwendung der algebraischen projektiven Geometrie.

Für die stabile Extraktion von signifikanten Punktmerkmalen werden zunächst verschiedene Interest-Operatoren untersucht. Der methodische Vergleich und die quantitative Analyse unterstreichen die Bedeutung einer robusten Gradientenberechnung. Zur Verbesserung des ausgewählten Lösungsansatzes werden Methoden für die Lokalisierung mit Subpixel-Genauigkeit, eine gleichmäßige Punktverteilung und die Nutzung von Farbinformation empfohlen.

Die Zuordnung dieser Merkmale stellt in Bildern, deren innere und äußere Orientierung unbekannt ist, eine besondere Herausforderung dar. Deshalb wird ein neues Verfahren zur merkmalsbasierten Korrespondenz-Analyse vorgestellt. Nach einem Verträglichkeitstest mit differentiellen Invarianten und der geometrischen Anpassung mit der Kleinste-Quadrate-Bildzuordnung wird die Ähnlichkeit durch die normierte Kreuz-Korrelation definiert. Mit einer hierarchischen Bildzuordnungs-Strategie wird die Erfolgsquote noch erheblich gesteigert, ohne dabei die Flexibilität des Ansatzes einzuschränken.

Die relative Orientierung der Bilder zueinander kann mit Hilfe der projektiven Geometrie auch ohne Kenntnis der inneren Orientierung bestimmt werden. Ein entscheidendes Problem bei der Automatisierung ist allerdings die robuste Erkennung und Beseitigung von falsch zugeordneten Merkmalen.

Daher wird ein neues Verfahren vorgeschlagen, das auf der Grundlage eines genetischen Algorithmus den üblichen Berechnungsaufwand signifikant reduziert. Anschließend können die einzelnen Bildpaare bzw. -tripel zu einem gemeinsamen Bildverband kombiniert werden, wobei die projektive Bündel-Ausgleichung eine spannungsfreie Rekonstruktion gewährleistet. Für die Auto-Kalibrierung im Nahbereich müssen fokussierbare Aufnahmesysteme modelliert werden, so dass ein Ansatz auf der Basis der dualen absoluten Fläche zweiter Ordnung verwendet wird. Zur Verbesserung der Ergebnisse werden besondere Eigenschaften der essentiellen Matrix ausgenutzt.

Die Rekonstruktion von Oberflächen kann erheblich vereinfacht und beschleunigt werden, wenn die konvergenten Bildpaare so umgerechnet werden, dass sie dem Stereo-Normalfall entsprechen. Um kritische Konfigurationen im Nahbereich zu vermeiden, wird ein hybrider Ansatz zur Normalbild-Erstellung beschrieben, der durch die Verwendung von Polarkoordinaten beliebige Bildanordnungen erlaubt. Als weiteres Anwendungsbeispiel wird eine volumetrische 3D-Rekonstruktion vorgestellt. Dabei wird durch effizientes Ausschneiden von Objektkonturen aus einem Volumen ein räumliches Modell erstellt.

In der vorliegenden Arbeit konnte nachgewiesen werden, dass mit den beschriebenen Verfahren eine automatische Auswertung von Bildern im Nahbereich auch ohne Zusatzinformation möglich ist. Die photogrammetrischen Methoden können dadurch in vielen Anwendungsbereichen weitaus flexibler und leistungsfähiger eingesetzt werden.

## Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Fachgebiet Photogrammetrie und Kartographie der Technischen Universität Berlin. Besonders danke ich Prof. JÖRG ALBERTZ und Prof. OLAF HELLWICH für die wissenschaftliche Betreuung und die wohlwollende Förderung der Arbeit. Bei Prof. REINHARD KLETTE und Prof. LOTHAR GRÜNDIG möchte ich mich für das gezeigte Interesse und die vielen hilfreichen Kommentare ganz herzlich bedanken.

In gleicher Weise danke ich allen Kollegen des Fachgebiets für die fruchtbare Zusammenarbeit. Besonders erwähnt seien YASEMIN KUZU, OLAF SINRAM und FRANK SCHOLTEN, die mir ihre Bilddaten zur Verfügung gestellt haben. Weiterhin führten viele konstruktive Diskussionen mit MATTHIAS HEMMLEB, ANDREAS REIGBER, ALBERT WIEDEMANN, MARTIN GRÜNDIG und TINO SCHUMACHER zu zahlreichen Ideen, die für meine Arbeit sehr wertvoll waren.

Nicht zuletzt gilt der Dank GERHARD KÖNIG und HEIKO SCHLICHTING, die mir immer mit Rat und Tat zur Seite standen sowie VERA HEINAU und STEPHAN GEHRKE, die mich bei den zeitaufwendigen Korrekturarbeiten unterstützten. Schließlich danke ich allen Freunden und Verwandten, die während der intensiven Promotionsarbeit häufig auf mich verzichten mussten, für ihr großes Verständnis und den geleisteten Beistand.

Diese Arbeit wurde zum Teil von der OTTO-LACMANN-Stiftung gefördert. Die methodischen Entwicklungen haben mir viel Spaß bereitet und ich hoffe, dass sie in den verschiedenen Anwendungsbereichen der Photogrammetrie und Computer Vision zum Einsatz kommen.

Berlin, im Januar 2004

VOLKER RODEHORST

*„In den Wissenschaften ist viel Gewisses,  
sobald man sich von den Ausnahmen  
nicht irre machen lässt und  
die Probleme zu ehren weiß.“*

JOHANN WOLFGANG V. GOETHE [1749-1832];  
dt. Dichter, Wilhelm Meisters Wanderjahre II

# Inhaltsverzeichnis

<b>1</b>	<b>EINLEITUNG UND EINORDNUNG .....</b>	<b>1</b>
1.1	THEMA UND PROBLEMSTELLUNG.....	1
1.2	ZIELSETZUNG UND METHODIK.....	2
1.3	SCHWERPUNKTE UND GLIEDERUNG .....	3
1.4	HISTORISCHE HINWEISE UND AUFNAHMETECHNIK .....	4
1.4.1	Entwicklung der Photogrammetrie.....	4
1.4.2	Entwicklung der Kameras .....	6
1.4.3	Anwendungsgebiete der Nahbereichs-Photogrammetrie .....	9
<b>2</b>	<b>MERKMALSEXTRAKTION .....</b>	<b>11</b>
2.1	GRUNDLAGEN DER BILDVERARBEITUNG .....	12
2.1.1	Digitale Bildrepräsentation .....	12
2.1.2	Lokale Operatoren.....	13
2.1.3	Glättungs-Filter .....	13
2.1.4	Differential-Operatoren.....	13
2.1.5	Implementations-Details .....	15
2.2	INTEREST-OPERATOREN .....	16
2.2.1	Anforderungen .....	17
2.2.2	Stand der Forschung.....	17
2.2.3	PLESSEY Punkt-Detektor .....	18
2.2.4	FÖRSTNER-Operator .....	19
2.2.5	Vergleich der Methoden.....	21
2.3	VERBESSERUNGEN UND ERWEITERUNGEN.....	25
2.3.1	Subpixel-Bestimmung durch Paraboloid-Anpassung .....	25
2.3.2	Gleichmäßige Punktverteilung mit adaptivem Schwellwert.....	29
2.3.3	Erweiterung der Autokorrelation für Farbbilder .....	30
2.4	EXPERIMENTELLE ERGEBNISSE .....	31
2.5	DISKUSSION .....	33

<b>3</b>	<b>MERKMALSBASIERTE KORRESPONDENZ-ANALYSE .....</b>	<b>35</b>
3.1	DEFINITION DER VERGLEICHSMUSTER .....	36
3.2	ÄHNLICHKEITSMASSE .....	37
3.2.1	Differenz-Korrelation .....	38
3.2.2	Normierte Kreuz-Korrelation .....	38
3.2.3	Differentielle Invarianten .....	39
3.2.4	Vergleich der Methoden .....	41
3.3	GEOMETRISCHE ANPASSUNG .....	42
3.3.1	Deformations-Modell .....	42
3.3.2	Kleinste-Quadrate-Bildzuordnung .....	44
3.3.3	Experimentelle Ergebnisse .....	47
3.4	VERFAHREN ZUR KORRESPONDENZ-ANALYSE .....	48
3.5	BILDZUORDNUNGS-STRATEGIEN .....	51
3.5.1	Annahmen und Einschränkungen .....	51
3.5.2	Hierarchischer Ansatz .....	52
3.6	EXPERIMENTELLE ERGEBNISSE .....	53
3.7	DISKUSSION .....	55
<b>4</b>	<b>PROJEKTIVE STEREO-GEOMETRIE .....</b>	<b>57</b>
4.1	GRUNDLAGEN DER PROJEKTIVEN GEOMETRIE .....	57
4.1.1	Hierarchie geometrischer Räume .....	58
4.1.2	Homogene Koordinaten .....	59
4.1.3	Dualität von Punkten und Linien .....	60
4.1.4	Geometrische Repräsentationen .....	61
4.1.5	Projektive Transformationen .....	64
4.2	GEOMETRISCHES KAMERAMODELL .....	68
4.2.1	Abbildungsmodell .....	69
4.2.2	Die Projektions-Matrix $P$ .....	69
4.2.3	Interpretation von $P$ .....	72
4.3	ORIENTIERUNG EINES BILDPAARES .....	75
4.3.1	Epipolar-Geometrie .....	76
4.3.2	Die Fundamental-Matrix $F$ .....	77
4.3.3	Normierter 8-Punkt-Algorithmus .....	79
4.3.4	Minimaler 7-Punkt-Algorithmus .....	80
4.3.5	Nicht-lineare Optimierung von $F$ .....	81
4.3.6	Degenerierte Konfigurationen .....	83
4.3.7	Die essentielle Matrix $E$ .....	84

---

4.4	ROBUSTE PARAMETERSCHÄTZUNG .....	85
4.4.1	Genetischer Algorithmus GASAC .....	85
4.4.2	Robustes Orientierungs-Verfahren.....	88
4.4.3	Experimentelle Ergebnisse .....	90
4.5	TRIANGULATION.....	91
4.5.1	Lineare Triangulation.....	92
4.5.2	Optimale Triangulation .....	92
4.5.3	Experimentelle Ergebnisse .....	95
4.6	DISKUSSION .....	96
<b>5</b>	<b>PROJEKTIVE MEHRBILD-GEOMETRIE.....</b>	<b>97</b>
5.1	ORIENTIERUNG EINES BILDTRIPELS .....	97
5.1.1	Trifokal-Geometrie.....	97
5.1.2	Minimaler 6-Punkt-Algorithmus.....	98
5.1.3	Robustes Orientierungs-Verfahren.....	100
5.1.4	Mehrbild-Tensoren.....	100
5.2	ANPASSEN DER PROJEKTIVEN BASIS .....	102
5.2.1	Räumliche Homographie.....	103
5.2.2	Projektive Faktorisierung .....	103
5.3	PROJEKTIVE BÜNDEL-AUSGLEICHUNG .....	105
5.4	EXPERIMENTELLE ERGEBNISSE .....	106
5.5	DISKUSSION .....	107
<b>6</b>	<b>AUTOMATISCHE SELBST-KALIBRIERUNG.....</b>	<b>109</b>
6.1	VERWENDUNG ZUSÄTZLICHER INFORMATIONEN .....	110
6.2	PRINZIP DER AUTO-KALIBRIERUNG .....	112
6.2.1	Affine Rekonstruktion mit der idealen Ebene.....	112
6.2.2	Innere Orientierung durch den absoluten Kegelschnitt.....	113
6.3	KLASSISCHE METHODEN ZUR AUTO-KALIBRIERUNG .....	114
6.3.1	Die KRUPPA-Gleichungen .....	115
6.3.2	Identische Singulärwerte der essentiellen Matrix .....	116
6.4	KALIBRIERUNG MIT DEM DUALEN ABSOLUTEN QUADRIC.....	117
6.4.1	Lineare Auto-Kalibrierung.....	117
6.4.2	Verbesserungen und Erweiterungen.....	119
6.4.3	Nicht-lineare Optimierung .....	120
6.4.4	Kombiniertes Verfahren .....	121
6.5	EXPERIMENTELLE ERGEBNISSE .....	121
6.6	DISKUSSION .....	123

---

<b>7</b>	<b>METRISCHE 3D-REKONSTRUKTION .....</b>	<b>125</b>
7.1	NORMALBILD-ERSTELLUNG .....	125
7.1.1	Lineare Bild-Transformation.....	126
7.1.2	Nicht-lineare Bild-Transformation.....	128
7.1.3	Experimentelle Ergebnisse .....	131
7.2	VOLUMETRISCHE 3D-MODELLIERUNG .....	133
7.2.1	Hierarchische Berechnung mit Octrees.....	134
7.2.2	Experimentelle Ergebnisse .....	135
7.3	DISKUSSION .....	136
<b>8</b>	<b>ZUSAMMENFASSUNG UND AUSBLICK .....</b>	<b>137</b>
	<b>ANHANG A.....</b>	<b>139</b>
A.1	SINGULÄRWERT-ZERLEGUNG .....	139
A.2	LINEARE KLEINSTE-QUADRATE-LÖSUNG .....	139
A.2.1	Homogene Gleichungssysteme .....	139
A.2.2	Inhomogene Gleichungssysteme.....	140
A.3	ROBUSTE SCHÄTZVERFAHREN.....	140
A.3.1	M-Schätzer (HUBER, TUKEY).....	141
A.3.2	Verfahren des Least-Median-of-Squares (LMedS).....	142
A.3.3	Monte-Carlo-Methode (RANSAC).....	143
	<b>LITERATURVERZEICHNIS .....</b>	<b>145</b>
	<b>FORMELZEICHEN .....</b>	<b>155</b>

# 1 Einleitung und Einordnung

## 1.1 Thema und Problemstellung

Die Photogrammetrie beschäftigt sich mit der Gewinnung geometrischer Objektinformationen durch die Auswertung von photographischen Bildern. Um die Objektgeometrie zu bestimmen, müssen die bei der Aufnahme geltenden Abbildungsgesetze rekonstruiert werden. Ein Punkt im dreidimensionalen Objektraum kann aber nur dann festgelegt werden, wenn er in mindestens zwei Bildern wiedergegeben ist und die Positionsunterschiede entsprechend ausgewertet werden (*Triangulations-Verfahren*).

Deshalb bestehen in der Photogrammetrie stets zwei Hauptaufgaben:

- **Korrespondenz-Analyse:** Zunächst müssen in den Bildern einander entsprechende (*homologe*) Objektpunkte identifiziert werden.
- **Orientierungs-Verfahren:** Durch die Rekonstruktion der geometrischen Beziehungen zwischen dem Objektraum und den Bildern können anschließend die Raumkoordinaten der zugehörigen Objektpunkte bestimmt werden.

Die Photogrammetrie ist ein berührungsloses (*indirektes*) Messverfahren, das im Gegensatz zu Laserscannern [WEH98] und Systemen mit strukturierter Beleuchtung [KLE96b] als *passives* Verfahren gilt. Das klassische Hauptanwendungsgebiet liegt in der topographischen Erfassung der Erdoberfläche mit Hilfe von Luftbildern (*Luftbild-Photogrammetrie*) und Satellitenaufnahmen (*Fernerkundung*) [ALB01]. Für diesen Bereich ist sowohl die Aufnahmetechnik mittels Reihemesskameras als auch die Auswertetechnik mit dem Einsatz von modernen Bildverarbeitungsmethoden weit entwickelt und in vieler Hinsicht ausgereift.

Im Nahbereich mit seinen stark variierenden Messaufgaben haben sich Aufnahme- und Auswertetechnik weit weniger einheitlich entwickelt. Durch den Einsatz fokussierbarer Kameras sowie die zunehmende Nutzung von Video- und CCD-Kameras werden immer neue Anwendungen möglich, bei denen aber die Abbildungsgeometrie nicht bekannt oder instabil ist. Diese Entwicklung trägt dazu bei, dass für den Nahbereich neuartige Bildorientierungsverfahren benötigt werden. Wegen der flexiblen Bildanordnungen und der perspektivischen Verzerrungen der Objekte in den Bildern werden allerdings besonders hohe Anforderungen an den Auswerteprozess gestellt.

Die photogrammetrischen Methoden stehen in enger Verwandtschaft sowohl zur Messtechnik als auch zum Bereich *Computer Vision*. Hier wurde seit etwa 1992 das Problem der Orientierung von Kameras mit Hilfe der *projektiven Geometrie* untersucht [MAY92, FAU92, HAR92b]. Die Verwendung der algebraischen Repräsentation erfordert zwar einen höheren Abstraktionsgrad als bei den klassischen geometrischen Methoden. Für die Automatisierung der Bildauswertung müssen die einzelnen Arbeitsschritte jedoch nicht unbedingt anschaulich sein.

Im Zuge einer automatischen Auswertung können kostenintensive interaktive Auswerteeinstrumente entfallen und durch handelsübliche Computer ersetzt werden. Wenn ein hoher Automatisierungsgrad erreicht wird, können auch nicht spezialisierte Nutzer eine photogrammetrische Aufnahme und Auswertung durchführen. Inzwischen hat das Verständnis der projektiven Geometrie einen fortgeschrittenen Stand erreicht [FAU01, HAR00, POL99c]. Obwohl die Theorie der linearen projektiven Beziehungen auch in der Photogrammetrie schon länger bekannt ist [BRA96], wird sie in der Praxis eher selten eingesetzt. Verstärkte Forschungsaktivitäten im Bereich der Photogrammetrie lassen aber vermuten, dass in Zukunft der Einsatz der projektiven Geometrie wesentlich an Bedeutung gewinnen wird [FÖR00, MAY02].

## 1.2 Zielsetzung und Methodik

Aufgabe der vorliegenden Arbeit ist es, die photogrammetrische 3D-Rekonstruktion im Nahbereich zu automatisieren und mit Hilfe einer Selbst-Kalibrierung flexibler zu gestalten. Dieses Ziel soll durch die Verbesserung und Integration von aktuellen Forschungsergebnissen in einen durchgängigen Auswerteprozess erreicht werden. Eine Übersicht der dazu erforderlichen Verarbeitungsschritte ist in Abb. 1.1 skizziert.

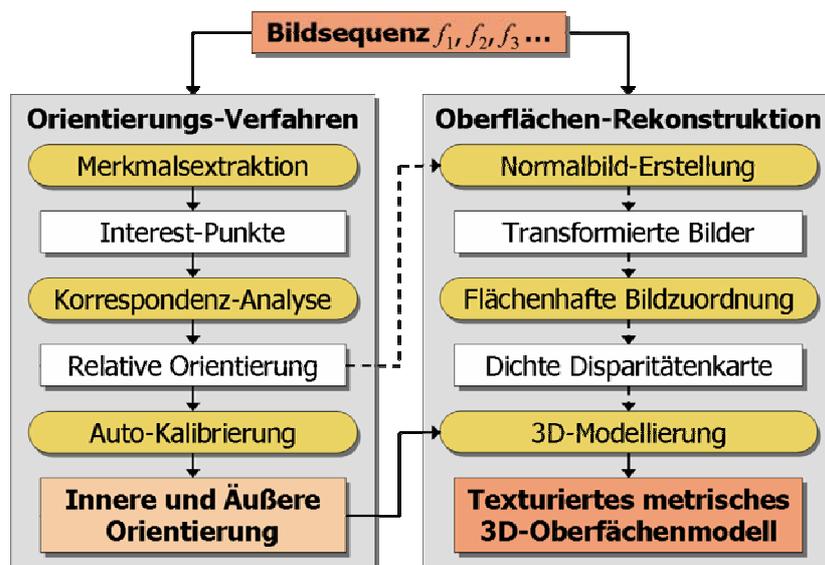


Abb. 1.1: Verarbeitungsschritte einer 3D-Rekonstruktion

Für das *Orientierungs-Verfahren* müssen zunächst mit Hilfe einer *Merkmalsextraktion* signifikante Bildmerkmale (z.B. Interest-Punkte) bestimmt werden. Anschließend müssen diese Punktmerkmale durch eine *Korrespondenz-Analyse* zugeordnet werden, so dass die Orientierung der Bilder zueinander bestimmt werden kann (Projektive relative Orientierung). Mit einer automatischen Selbst-Kalibrierung (*Auto-Kalibrierung*) kann schließlich die gesuchte innere und äußere Orientierung berechnet und die zugrunde liegende Aufnahmegeometrie rekonstruiert werden.

Die Beschränkung auf signifikante Bildelemente liefert allerdings nur einzelne Objektpunkte. Für eine zusammenhängende *Oberflächen-Rekonstruktion* wird im Allgemeinen eine *flächenhafte Bildzuordnung* eingesetzt. Dieser Vorgang kann erheblich vereinfacht und beschleunigt werden, wenn zuvor durch eine geometrische Transformation *Normalbilder* erstellt werden. Nach der *3D-Modellierung* erhält man schließlich eine geometrische Rekonstruktion der abgebildeten Objekte.

### 1.3 Schwerpunkte und Gliederung

Die Schwerpunkte dieser Arbeit liegen in der Entwicklung einer merkmalsbasierten Korrespondenz-Analyse und einer optimierten Auto-Kalibrierung mit Hilfe der algebraischen projektiven Geometrie. Aufgrund der vielfältigen Teilprobleme orientiert sich die Gliederung der vorliegenden Arbeit an der Reihenfolge der Verarbeitungsschritte.

- Zunächst werden für die *Merkmalsextraktion* in *Kapitel 2* verschiedene Interest-Operatoren bewertet. Um die klassischen Lösungsansätze zu verbessern, werden Methoden für eine Lokalisierung mit Subpixel-Genauigkeit, eine gleichmäßige Punktverteilung und die Nutzung von Farbinformationen vorgestellt.
- In *Kapitel 3* wird durch die zweckmäßige Kombination bekannter Korrelations-Methoden ein neues Verfahren zur merkmalsbasierten *Korrespondenz-Analyse* vorgeschlagen. In Verbindung mit einer hierarchischen Bildzuordnungs-Strategie wird zuerst ein Verträglichkeitstest mit differentiellen Invarianten durchgeführt. Nach der geometrischen Anpassung der Merkmale mit der Kleinste-Quadrate-Bildzuordnung wird deren Ähnlichkeit schließlich durch die normierte Kreuz-Korrelation bestimmt.
- Das *Kapitel 4* behandelt die *relative Orientierung* eines Bildpaares mit Hilfe der projektiven Geometrie. Für die automatische Erkennung und Beseitigung von falsch zugeordneten Merkmalen wird ein neues Verfahren zur *robusten Parameterschätzung* vorgestellt, das auf der Grundlage eines genetischen Algorithmus den üblichen Berechnungsaufwand signifikant reduziert. Anschließend wird eine Methode zur Bestimmung von räumlichen Koordinaten im projektiven Raum durch *Triangulation* erläutert.

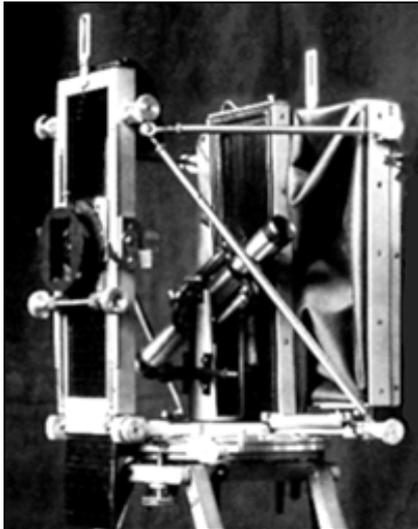
- Um die einzelnen projektiven Rekonstruktionen in einem *gemeinsamen Bildverband* zu kombinieren, werden in *Kapitel 5* zunächst verschiedene Näherungsverfahren analysiert. Danach wird durch den Einsatz einer globalen projektiven Bündel-Ausgleichung eine spannungsfreie Aufnahmegeometrie erzeugt.
- Das *Kapitel 6* stellt den Ansatz zur *automatischen Selbst-Kalibrierung* vor, der für die Anwendung im Nahbereich auch fokussierbare Aufnahmesysteme unterstützt. Dafür wird ein Verfahren auf der Basis der dualen absoluten Fläche zweiter Ordnung verwendet und die Ergebnisse durch Eigenschaften der essentiellen Matrix verbessert.
- Nach der Rekonstruktion der Aufnahmegeometrie befasst sich das *Kapitel 7* mit der *3D-Modellierung* der abgebildeten Objekte. Um den Aufwand zur Erstellung von Oberflächenmodellen zu reduzieren, wird ein hybrider Ansatz zur *Normalbild-Erstellung* beschrieben, der durch die Verwendung von Polarkoordinaten beliebige Bildanordnungen im Nahbereich erlaubt. Als Anwendungsbeispiel werden Ergebnisse einer volumetrischen 3D-Rekonstruktion durch Ausschneiden von Objektkonturen aus einem Volumen präsentiert.
- Das *Kapitel 8* enthält schließlich eine Zusammenfassung dieser Arbeit und gibt einen Ausblick auf zukünftige Entwicklungen.

## 1.4 Historische Hinweise und Aufnahmetechnik

### 1.4.1 Entwicklung der Photogrammetrie

Die Photogrammetrie geht auf das Jahr 1858 zurück. Damals hatte der Architekt MEYDENBAUER die Idee, das häufig riskante Handaufmass für die Dokumentation von wichtigen Architekturdenkmälern zu umgehen und photographische Bilder für die Fassadenvermessung einzusetzen. Danach entwarf er eine Reihe von Messkameras, die aufgrund ihrer stabilen inneren Orientierung die Rekonstruktion des Aufnahmestrahlenbündels erlauben.

Im Messbild-Archiv beim Brandenburgischen Landesamt für Denkmalpflege sind noch etwa 16.000 Diapositive dieser Kameras auf Glasplatten in der Größe  $40 \times 40 \text{ cm}^2$  vorhanden (vgl. Abb. 1.2). Da die Kalibrierungsdaten von diesen Bildern verloren gegangen sind, muss die innere Orientierung teilweise aufwändig wiederhergestellt werden, wie z.B. bei der Rekonstruktion des Berliner Stadtschlusses [WIE00] oder dem Kloster CHORIN [ALB97]. Die in der vorliegenden Arbeit beschriebenen Methoden könnten die Auswertung dieser historischen Messbilder erheblich vereinfachen.



a) MEYDENBAUER-Kamera um 1890 [ALB97]



b) Messbild vom Berliner Stadtschloss

**Abb. 1.2:** Beispiel für das Messbild einer MEYDENBAUER-Kamera

Die weitere Entwicklung der Photogrammetrie war kurz nach 1900 stark durch die Einführung des stereoskopischen Messprinzips geprägt. Mit geeigneten Hilfsmitteln wurde es möglich, dass ein Beobachter durch stereoskopische Betrachtung schnell und sicher homologe Objektpunkte in zwei Bildern bestimmen kann. Wenig später begann die Entwicklung stereophotogrammetrischer Auswertegeräte, die die gesamte Abbildungsgeometrie mit *analogen* (optischen oder mechanischen) Mitteln rekonstruieren. Damit waren die beiden Hauptaufgaben zweckmäßig gelöst und die Photogrammetrie begann ihren Siegeszug, insbesondere bei der topographischen Kartierung aus Luftbildern.

Mit dem Aufkommen der elektronischen Datenverarbeitung ergaben sich neue methodische Möglichkeiten. Die Auswertetechnik ging zur *analytischen* Photogrammetrie über, in der die Abbildungszusammenhänge nach den Gesetzen der analytischen Geometrie rechnerisch behandelt werden. Inzwischen ist mit den Fortschritten in der digitalen Bildverarbeitung der Übergang zur *digitalen* Photogrammetrie eingeleitet worden. Ausgehend von digitalen Bilddaten kann die stereoskopische Punktbestimmung nun durch eine rechnerische Bildzuordnung ersetzt werden, wodurch ein hoher Automatisierungsgrad erreicht wird. Im Wesentlichen nutzt man in der digitalen Photogrammetrie aber weiterhin die Regeln der analytischen Geometrie.

In zunehmendem Maße setzen sich jedoch Bemühungen durch, die Zusammenhänge zwischen dem Objektraum und den von ihm gewonnenen Bildern mittels *projektiver Geometrie* in allgemeiner Form zu beschreiben. Damit ist auch eine Verallgemeinerung der Photogrammetrie verbunden, da von der traditionellen Form der Messkamera abstrahiert wird. Mit dieser Arbeit soll ein Beitrag zu dieser Entwicklung geleistet werden.



a) ZEISS TMK 6



b) Bildausschnitt mit Rahmenmarke

Abb. 1.3: Beispiel für das Bild einer terrestrischen Messkamera

## 1.4.2 Entwicklung der Kameras

### 1.4.2.1 Messkameras

Die klassische Messkamera ist eine speziell für photogrammetrische Anwendungen entwickelte Kamera mit stabiler und konstanter innerer Orientierung. Durch die Definition des Bildkoordinatensystems mit Rahmenmarken, die Verwendung von verzeichnungsfreien Hochleistungsobjektiven und die Verebnung der Bildfläche (z.B. durch Glasplatten als Emulsionsträger bzw. mechanische Andruck- oder Ansaugvorrichtungen für das Filmmaterial) kann eine hohe Genauigkeit erreicht werden.

Diese aufwändige Aufnahmetechnik wird in der Luftbildaufnahme seit Jahrzehnten in großem Umfang eingesetzt und in diesem Bereich auch in absehbarer Zukunft noch ihre Bedeutung behalten. Im Nahbereich werden konventionelle Messkameras hauptsächlich zur Aufnahme von statischen Objekten eingesetzt (vgl. Abb. 1.3). Da sie aber in der praktischen Anwendung wenig flexibel sind, wurde schon seit langem nach anderen Lösungen gesucht.

### 1.4.2.2 Teil-Messkameras

Bei der Teil-Messkamera handelt es sich im Prinzip um einen handelsüblichen Photoapparat im professionellen Mittelformat mit entsprechendem Bedienungskomfort (vgl. Abb. 1.4). Durch eine zusätzliche Glasplatte mit Réseaukreuzen vor dem Filmmaterial wird ein festes Bezugssystem für die Bildebene definiert. Die geometrischen Auswirkungen von Filmunebenheiten und Filmverzug können anschließend numerisch durch eine geeignete Transformation kompensiert werden.



a) ROLLEIFLEX 6006 metric



b) Bildausschnitt mit Réseaukreuzen

**Abb. 1.4:** Beispiel für das Bild einer Teil-Messkamera

Allerdings bewirkt die Glasgitterplatte leichte Unschärfe und zusätzliche Verzeichnungen. Da die abgebildeten Réseaukreuze bei einer automatischen Bildauswertung stören, sollten sie nach der Durchführung der Orientierung durch modellbasierte Mustererkennung entfernt werden.

### 1.4.2.3 Amateur-Kameras

Unter Amateur- oder Nichtmesskameras versteht man handelsübliche Digital-Kameras und Kleinbild-Photoapparate sowie Digital-Video-Camcorder oder Videokameras (vgl. Abb. 1.5). Für diese Aufnahmesysteme ist die innere Orientierung nicht oder nur ungenau bekannt. Diese vergleichsweise preiswerten Aufnahmesysteme zeichnen sich durch vielfältige Verwendungsmöglichkeiten aus und erlauben aufgrund des geringen Gewichts auch flexible Aufnahmen aus der Hand in beliebiger Stellung.

Durch die Verfügbarkeit einer Vielzahl an Objektiven (z.B. Zoom, Weitwinkel, Makro) und einstellbare Fokussierung können variable Objektgrößen erfasst werden. Einige Digital- und APS-Kameras zeichnen neben den Bildern auch zusätzliche Aufnahmedaten wie z.B. die verwendete Brennweite mit auf. Bei dynamischen Objekten können mit Camcordern oder durch Photoapparate mit einem motorisierten Bildtransport schnelle Bildfolgen erzeugt werden.

Neben der Tatsache, dass für derartige Kameras nur selten die innere Orientierung gegeben ist, weisen die Objektive mit variabler Brennweite häufig große und veränderliche Verzeichnungen auf. Deshalb muss eine besondere Kalibrierung durchgeführt werden. Eine umfassende Darstellung aktueller Aufnahmetechniken in der Nahbereichs-Photogrammetrie ist z.B. bei LUHMANN [LUH00] zu finden.



a) KODAK DCS-Pro 14n



b) SONY DCR-TRV 30E



c) CANON EOS IX

**Abb. 1.5:** Verwendete Digital-Kamera, DV-Camcorder und APS Spiegelreflex-Kamera

#### 1.4.2.4 Digitalisierung

Alle neuen photogrammetrischen Entwicklungen gehen von digitalen Bilddaten aus. Zur Digitalisierung der in der Photogrammetrie üblichen großformatigen Luftbilder ( $23 \times 23 \text{ cm}^2$ ) werden in der Regel spezielle Scanner eingesetzt. Für den Nahbereich sind die Möglichkeiten zur Erzeugung und Beschaffung von digitalem Bildmaterial dagegen sehr vielfältig geworden und haben sich in den letzten Jahren stark vereinfacht.

Digital-Kameras und DV-Camcorder speichern ihre Bilddaten derzeit auf ultrakompakten Speichermedien (z.B. CompactFlash, MicroDrive, Memory-Sticks, SmartMedia-, MultiMedia- oder Secure-Digital Cards) und können direkt bzw. über einen Adapter (z.B. USB oder IEEE 1394 Firewire) mit einem Auswerterechner verbunden werden. Bei analogen Videokameras kann eine elektronische Abtastung mit Videokarten (*Frame-Grabbern*) erfolgen, die teilweise schon zur Standardausrüstung gehören.

Falls Diapositive, Negativfilme oder Papierabzüge vorliegen, kann eine mechanische Abtastung z.B. mit Film-, Trommel- oder Flachbett-Scannern erfolgen. Viele Photolabors bieten bei der Entwicklung von Mittelformat-, APS- oder Kleinbild-Filmen bereits eine Photo-CD mit digitalisierten Abzügen an.

Kamera-Typ	Modell	Aufnahmematerial	Größe [mm]	Digitalisierung	Auflösung [dpi]	Bildgröße [Pixel]
MEYDENBAUER	Unbekannt	Glasplatte (sw)	400 × 400	Trommel-Scanner	1.080	17.000 × 17.000
Messkamera	ZEISS TMK 6	Glasplatte (sw)	100 × 80	Trommel-Scanner	2.540	10.000 × 8.000
Digital-Kamera	KODAK DCS-Pro 14n	CMOS-Sensor	36 × 24	-	-	4.536 × 3.024
Teil-Messkamera	ROLLEI 6006 metric	Rollfilm	55 × 55	Flachbett-Scanner	1.400	3.024 × 3.024
Spiegelreflex	CANON EOS-IX	APS-Rollfilm	30,2 × 16,7	Film-Scanner	1.800	2.137 × 1.240
DV-Camcorder	SONY DCR-TRV 30E	CCD-Sensor	3,2 × 2,4	-	-	1.360 × 1.020
Videokamera	SONY DXC-730P	CCD-Sensor	8,8 × 6,6	Frame-Grabber	-	768 × 581

**Tab. 1.1:** Beispiele für verschiedene Aufnahmetechniken zur Erzeugung von digitalem Bildmaterial mit den daraus resultierenden Bildgrößen

### 1.4.3 Anwendungsgebiete der Nahbereichs-Photogrammetrie

Photogrammetrische Verfahren können prinzipiell in allen Bereichen eingesetzt werden, in denen ein Messobjekt photographisch oder mit anderen technischen Mitteln abgebildet werden kann. Für die Auswertung müssen dann geeignete Verfahren eingesetzt werden, die u.a. die jeweilige Abbildungsgeometrie rekonstruieren. Beispiele für praktische Anwendungen der Nahbereichs-Photogrammetrie sind nach LUHMANN [LUH02]:

- **Architektur** (Archäologie und Denkmalpflege)
- **Ingenieurbau** (Hoch-, Tief- und Bergbau sowie Tunnelvermessung)
- **Industrielle Messtechnik** (Automobil-, Maschinen- und Schiffbau, Luft- und Raumfahrt, Robotik, Design und Reverse Engineering sowie Anlagenbau)
- **Medizin** (Navigation und Oberflächenerfassung)
- **Naturwissenschaften** (Strömungsmesstechnik, Wellentopographie, Geologie und Biologie)
- **Sonderanwendungen** (Materialprüfung, Forstwissenschaft, Unfall- und Tatortvermessung)

Die derzeitigen methodischen Entwicklungen, zu denen auch die vorliegende Arbeit beitragen soll, machen die photogrammetrischen Methoden flexibler bzw. leistungsfähiger und werden diese Anwendungsvielfalt noch weiter erhöhen.

*„Wer bei Kleinigkeiten keine Geduld hat,  
dem misslingt der große Plan.“*

KONFUZIUS [551-479 v. Chr.];  
chinesischer Philosoph

## 2 Merkmalsextraktion

Die Aufgabe der Merkmalsextraktion ist es, in den Bildern charakteristische Muster zu suchen, von denen angenommen werden kann, dass sie in möglichst vielen Bildern enthalten sind und sich daher als Kandidaten für eine Bildzuordnung eignen. In der Nahbereichs-Photogrammetrie werden dazu häufig spezielle Passpunkte mit eindeutigen Markierungen signalisiert, die in den Bildern leicht wieder erkannt werden können. In der Praxis existieren allerdings viele Situationen, in denen es sehr schwierig oder gar unmöglich ist, diese Markierungen geeignet anzubringen. Schließlich ist es für die Automatisierung des Messvorgangs vorteilhaft, wenn die gesuchten Merkmale ohne vorherige Präparierung der Objekte direkt aus den Bildinhalten abgeleitet werden können. Die Auto-Kalibrierung in Kapitel 6 verwendet daher nur das Bildmaterial und setzt keine zusätzliche durch Signalisierung erzeugte Passpunkt-Information voraus.

Signifikante Bildmerkmale zeichnen sich durch auffällige Intensitätsänderungen aus und können daher mit Hilfe des Bildgradienten lokalisiert werden. Da die Realisierung des Verfahrens einen entscheidenden Einfluss auf die Qualität der Merkmalsextraktion hat, wird nach einer kurzen Einführung in die digitale Bildverarbeitung zunächst auf die robuste Berechnung des Gradienten eingegangen (Abschnitt 2.1). Allerdings sind nicht alle geometrischen Merkmale gleichermaßen für die Bestimmung der Orientierung geeignet. So kann z.B. für eine vertikale Kante bei lokaler Betrachtung nur die horizontale Verschiebung bestimmt werden, wobei die tatsächliche Lage der Kante mehrdeutig bleibt. MALLOT [MAL98] bezeichnet dies als *Aperturproblem*.

In der Praxis werden deshalb häufig punktförmige Merkmale gesucht, die sich deutlich von ihrer Umgebung abheben. Für die Extraktion dieser markanten Bildpunkte kommen so genannte *Interest-Operatoren* zum Einsatz (Abschnitt 2.2). Durch einen umfassenden Vergleich führender Interest-Operatoren mit standardisierten Referenzbildern soll die leistungsfähigste Methode unter den Anforderungen der Bildzuordnung herausgefunden werden. Im Anschluss werden zu der klassischen Vorgehensweise noch diverse Verbesserungen vorgeschlagen (Abschnitt 2.3). Neben der Möglichkeit zur Lokalisierung mit Subpixel-Genauigkeit und der Optimierung der Punktverteilung wird auch eine Erweiterung für Farbbilder empfohlen. Abschließend werden die Ergebnisse der Merkmalsextraktion anhand realer Bildbeispiele illustriert (Abschnitt 2.4).

## 2.1 Grundlagen der Bildverarbeitung

Zu Beginn sollen einige grundlegende Definitionen und Annahmen erläutert werden. Die folgenden Abschnitte konzentrieren sich dabei im Wesentlichen auf die robuste Berechnung des Gradienten mit *kontinuierlichen Differential-Operatoren* (Abschnitt 2.1.4). Umfassende Informationen zu den Grundlagen der digitalen Bildverarbeitung sind z.B. in den Lehrbüchern von RUSS [RUS02], GONZALEZ und WOODS [GON01], JÄHNE [JÄH97] oder CASTLEMAN [CAS96] zu finden.

### 2.1.1 Digitale Bildrepräsentation

In der digitalen Bildverarbeitung wird ein Bild örtlich *diskretisiert*, d.h. als Raster mit kleinen flächenhaften Bildelementen (*Pixeln*) dargestellt. Der Ursprung des Bildrasters mit  $M \times N$  Pixeln wird in dieser Arbeit in der linken oberen Ecke angenommen, d.h. die Spalten werden von links nach rechts und die Zeilen von oben nach unten aufsteigend nummeriert. Die Ausrichtung des kartesischen Koordinatensystems wird dadurch als *rechtsdrehend* definiert und ist im mathematischen Sinn negativ orientiert (vgl. Abb. 2.1). Die Skalierung der Koordinatenachsen unterscheidet sich häufig in horizontaler und vertikaler Richtung, so dass die Pixel rechteckig und nicht quadratisch sind. Durch das Orthogonalraster wird eine Nachbarschaft definiert, die für die verschiedenen Verarbeitungsoperationen von großer Bedeutung ist.

Die *Bildfunktion*  $f(x, y)$  definiert einen bestimmten Intensitätswert an der Stelle der Ortskoordinaten  $(x, y)$  im Raster. Die Intensitäten werden in der Regel mit acht Bit *quantisiert*, so dass  $2^8 = 256$  Graustufen zwischen 0 (*schwarz*) und 255 (*weiß*) verfügbar sind. Echtfarbbilder  $f = (r, g, b)^T$  werden im Allgemeinen mit acht Bit je Farbkanal (*rot*, *grün* und *blau*) abgespeichert, so dass insgesamt  $2^{24} = 16,8$  Mio. Farbtöne definiert werden können. Neben dieser technisch motivierten Festlegung können aber auch weitere Farbmodelle verwendet werden [KOS98].

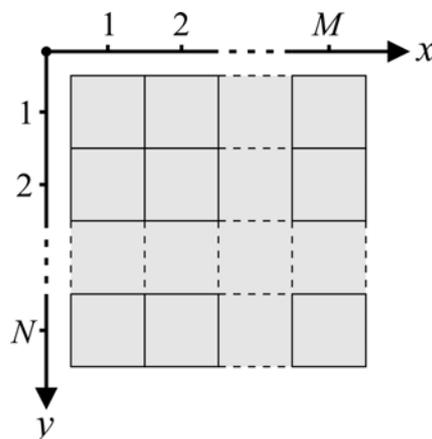


Abb. 2.1: Definition des Pixel-Koordinatensystems

## 2.1.2 Lokale Operatoren

Ein Standardverfahren zur Filterung eines Bildes mit Hilfe von linearen, verschiebungsinvarianten Operatoren ist die *Faltung*. Der Einsatz solcher Filteroperatoren in der Bildverarbeitung ist in KLETTE und ZAMPERONI [KLE96a] sowie RADIG [RAD93] umfassend dargelegt. Bei der Anwendung von lokalen Filtern wird eine quadratische Maske mit dem Radius  $r$  über das Bild  $f$  verschoben und an jeder Position mit

$$f(x, y) * g(x, y) = \sum_{i=-r}^r \sum_{j=-r}^r f(x-i, y-j) \cdot g(i, j) \quad (2.1)$$

gefaltet, wobei das Ergebnis durch den *Filterkern*  $g$  gesteuert wird. Es ist zu beachten, dass diese Formulierung der diskreten Faltung im mathematischen Sinn erfolgt, wobei der Bildinhalt in gespiegelter Form multipliziert wird. Diese Eigenschaft ist für das korrekte Vorzeichen bei unsymmetrischen Filtermasken wichtig, wie z.B. den Ableitungsfiltern 1. Ordnung (Abschnitt 2.1.4).

## 2.1.3 Glättungs-Filter

Die Verwendung eines Glättungs-Filters (*Tiefpass*) vermindert das Rauschen und definiert den Maßstab der Merkmale, die erkannt werden sollen. Gebräuchlich ist die Verwendung der zweidimensionalen GAUSSschen Normalverteilung

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad (2.2)$$

wobei  $\sigma$  die *Standardabweichung* und  $\sigma^2$  die *Varianz* bezeichnen (siehe Abb. 2.3a). Bei der Festlegung des Wirkungsbereichs durch die Standardabweichung  $\sigma$  ist stets ein Kompromiss zwischen Rauschunterdrückung und Detailverlust zu treffen.

## 2.1.4 Differential-Operatoren

Nachfolgend sollen Differential-Operatoren beschrieben werden, die einer Funktion ihre Ableitung zuordnen. Sie werden eingesetzt, um den Betrag und die Richtung von Grauwert-Änderungen in einem Bild zu bestimmen. Im Gegensatz zur groben Approximation durch direkte Grauwert-Differenzen oder zum Einsatz von diskreten  $3 \times 3$ -Masken, z.B. dem klassischen SOBEL-Operator, werden in dieser Arbeit parametrisierbare kontinuierliche Filter verwendet.

### 2.1.4.1 Ableitungsfiler 1. Ordnung

CANNY [CAN86] hat eine kontinuierliche Filterfunktion für Stufenkanten beschrieben, die mit Hilfe der ersten Ableitung der GAUSS-Funktion  $G$  approximiert wird. Die Faltung und Differentiation sind lineare verschiebungsinvariante Operationen, für die Assoziativ- und Kommutativgesetze gelten.

$$\underbrace{\text{Gestörtes Bild} * \text{GAUSS-Funktion}}_{\text{Geglättetes Bild}} * \text{Differentiation} = \text{Gestörtes Bild} * \underbrace{\text{GAUSS-Funktion} * \text{Differentiation}}_{\text{Ableitung der GAUSS-Funktion}}$$

**Abb. 2.2:** Zusammenhang von Differentiation und Glättung

Die Ableitung eines geglätteten Bildes entspricht daher einer Faltung des Bildes mit der Ableitung der GAUSS-Funktion. Diese grundlegende Eigenschaft ist in Abb. 2.2 dargestellt.

Für die Bestimmung der partiellen Ableitungen einer zweidimensionalen Bildfunktion  $f$  in  $x$ - und  $y$ -Richtung müssen zunächst die ersten Ableitungen der GAUSS-Funktion (vgl. Abb. 2.3) in Abhängigkeit von der Standardabweichung  $\sigma$  bestimmt

$$\begin{aligned} G_x(x, y) &= \frac{\partial G(x, y)}{\partial x} = -\frac{x}{2\pi\sigma^4} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \\ G_y(x, y) &= \frac{\partial G(x, y)}{\partial y} = -\frac{y}{2\pi\sigma^4} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \end{aligned} \quad (2.3)$$

und anschließend als kontinuierlicher Filterkern mit dem Bild gefaltet werden

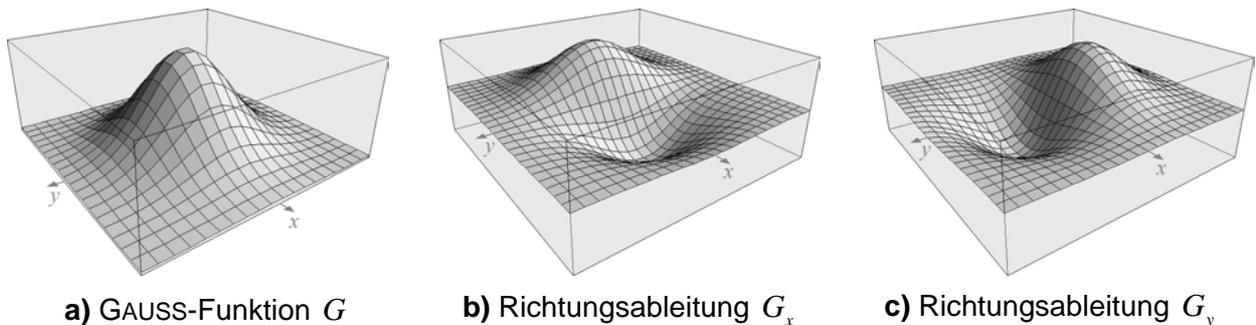
$$\begin{aligned} f_x(x, y) &= \frac{\partial(f(x, y) * G(x, y))}{\partial x} = f(x, y) * G_x(x, y), \\ f_y(x, y) &= \frac{\partial(f(x, y) * G(x, y))}{\partial y} = f(x, y) * G_y(x, y). \end{aligned} \quad (2.4)$$

Ein Beispiel für die partielle Ableitung ist in Abb. 2.5c und d dargestellt. Die einfachste richtungsabhängige, lineare Kombination von differentiellen Operatoren bezeichnet man als *Intensitätsgradient*

$$\nabla f(x, y) = [f_x(x, y), f_y(x, y)]^T, \quad (2.5)$$

dessen Betrag folgendermaßen definiert ist:

$$|\nabla f(x, y)| = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad (2.6)$$



**Abb. 2.3:** Beispiel für eine GAUSS-Funktion mit  $\sigma = 1,4$  und ihre beiden Richtungsableitungen

### 2.1.4.2 Ableitungsfiler 2. Ordnung

Sprunghafte Änderungen der Bildfunktion werden nicht nur durch die Extrema der ersten Ableitung charakterisiert, sondern können auch durch die Nullstellen der zweiten Ableitung erkannt werden. In dieser Arbeit wird der Differential-Operator allerdings nicht zur Merkmalsextraktion, sondern für die Paraboloid-Anpassung in Abschnitt 2.3.1 benötigt.

Um die zweite Ableitung einer Bildfunktion zu bestimmen, müssen zunächst die partiellen Ableitungen der GAUSS-Funktion zweiter Ordnung

$$\begin{aligned} G_{xx}(x, y) &= \frac{\partial^2 G(x, y)}{\partial x^2} = \frac{x^2 - \sigma^2}{2\pi\sigma^6} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \\ G_{yy}(x, y) &= \frac{\partial^2 G(x, y)}{\partial y^2} = \frac{y^2 - \sigma^2}{2\pi\sigma^6} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \end{aligned} \quad (2.7)$$

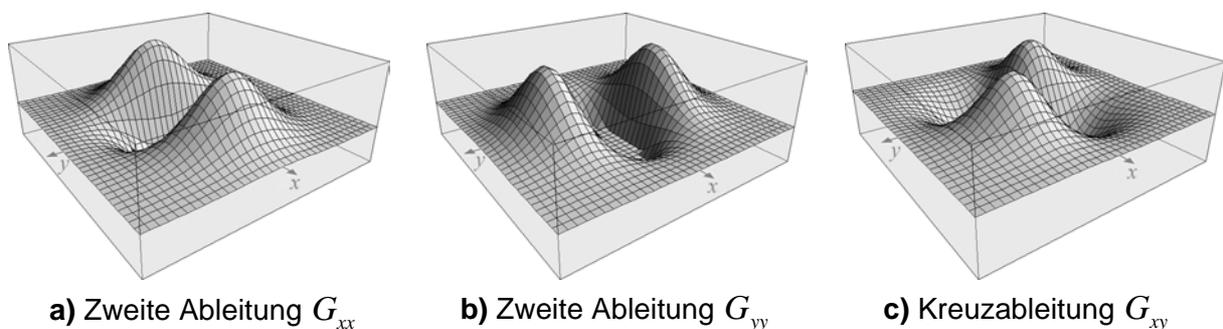
oder die Kreuzableitung

$$G_{xy}(x, y) = \frac{\partial^2 G(x, y)}{\partial x \partial y} = \frac{xy}{2\pi\sigma^6} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2.8)$$

mit der Standardabweichung  $\sigma$  als Parameter bestimmt werden (siehe Abb. 2.4). Anschließend muss das Bild  $f$  mit dem entsprechenden kontinuierlichen Filterkern gefaltet werden.

### 2.1.5 Implementations-Details

Bei der Faltung muss je nach Einflussbereich des Filterkerns auf Bereiche außerhalb des Bildes zugegriffen werden (vgl. Gleichung 2.1). Diese *Randprobleme* können beseitigt werden, indem die Bildgrenzen an jeder Seite um den Maskenradius  $r$  erweitert werden und die Intensitätswerte z.B. durch Spiegelung am Bildrand ergänzt werden. Der Berechnungsaufwand zur Bestimmung der Ableitungen hängt von der *Dimension* des Faltungskerns ab. Der Radius  $r$  der verwendeten Filtermasken kann in Abhängigkeit von der Standardabweichung  $\sigma$  mit  $r = \sqrt{2}\pi\sigma$  automatisch geschätzt werden.



**Abb. 2.4:** Die partiellen Ableitungen der GAUSS-Funktion 2. Ordnung und ihre Kreuzableitung

Die kontinuierlichen Differential-Operatoren werden also nur durch den Parameter  $\sigma$  in einem typischen Bereich von 0,5 bis 3,0 gesteuert, wodurch der Glättungseffekt beeinflusst wird. Durch eine lokale Statistik kann auch dieser Parameter geschätzt werden [STE98].

Die von DERICHE [DER90] vorgeschlagene Implementierung von CANNYS Ansatz gewährleistet durch einen festen Faltungskern einen konstanten Berechnungsaufwand. Die Auswahl von bestimmten Einflussgebieten wird beim DERICHE-Operator mit Hilfe von rekursiven Bilddurchläufen erreicht. Eine Erweiterung des CANNYSchen Qualitätskriteriums [CAN86] um die benachbarten Kantenpunkte zeigt jedoch, dass die Impulsantwort der ersten Ableitung einer GAUSS-Funktion der Lokalisierungsgüte von DERICHE überlegen ist [TAR90].

Des Weiteren wurde bei der Realisierung die *Separierbarkeit* von Operatoren ausgenutzt. Unter Verwendung der Eigenschaft

$$\exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) = \exp\left(-\frac{x^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{y^2}{2\sigma^2}\right) \quad (3.1)$$

kann eine Zerlegung der zweidimensionalen Filterung in zwei eindimensionale Filterungen vorgenommen werden. Durch eine zeilenweise Prozessierung, gefolgt von einer spaltenweisen Bearbeitung, kann der quadratische Berechnungsaufwand  $O(N^2)$  auf ein lineares Maß  $O(N)$  reduziert werden.

## 2.2 Interest-Operatoren

Als *Interest-Operatoren* bezeichnet man Verfahren zur Extraktion markanter Bildstellen, die in ihrer Umgebung möglichst einzigartig sind und mit hoher Wahrscheinlichkeit im korrespondierenden Bild in ähnlicher Weise abgebildet werden. Gleichzeitig liefern die Operatoren eine oder mehrere Kenngrößen (*Interest-Werte*), die bei der späteren Bildzuordnung verwendet werden können. Die Bedeutung der extrahierten Merkmale ist allerdings abhängig vom Kontext und daher entsprechen die Interest-Punkte nicht notwendigerweise physikalischen Ecken in der Szene.

Ausgehend von den Kriterien eines optimalen Interest-Operators soll aus der Vielzahl existierender Methoden zunächst eine Vorauswahl getroffen werden (Abschnitt 2.2.2). Danach wird auf der Grundlage der robusten Gradientenberechnung die Implementierung des PLESSEY Punkt-Detektors (Abschnitt 2.2.3) und des FÖRSTNER-Operators (Abschnitt 2.2.4) beschrieben. Schließlich soll durch einen umfassenden Vergleich mit standardisierten Referenzbildern der leistungsfähigste Interest-Operator für die Bildzuordnung bestimmt werden (Abschnitt 2.2.5).

### 2.2.1 Anforderungen

Zunächst sollen die Anforderungen an einen optimalen Interest-Operator definiert werden. Als Kriterien für einen markanten Zuordnungskandidaten können nach HARALICK und SHAPIRO [HAR92a] sowie LUHMANN [LUH00] folgende Merkmale dienen:

- **Deutlichkeit:** Ein Interest-Punkt sollte sich deutlich vom Hintergrund abheben und aufgrund der charakteristischen Eigenschaften in seiner Umgebung einzigartig sein.
- **Invarianz:** Die Bestimmung eines Interest-Punktes sollte unabhängig von den zu erwartenden geometrischen und radiometrischen Verzerrungen sein.
- **Robustheit:** Die Auswahl der Interest-Punkte sollte gegenüber Rauschen und groben Fehlern wenig empfindlich sein.
- **Seltenheit:** Neben der lokalen Deutlichkeit sollte ein Interest-Punkt auch eine globale Einzigartigkeit besitzen, um z.B. bei der Korrespondenz-Analyse (Kapitel 3) die Unterscheidbarkeit von sich wiederholenden Mustern zu verbessern.
- **Interpretierbarkeit:** Interest-Werte sollten eine sinnvolle Bedeutung haben, damit sie zur Korrespondenz-Analyse und höheren Bildinterpretation herangezogen werden können.

Während die Eigenschaften der Deutlichkeit, Invarianz und Robustheit die wesentlichen Anforderungen an einen Interest-Operator definieren, verschärfen die Eigenschaften der Seltenheit und Interpretierbarkeit die Bedeutung des Begriffs ‚*interessant*‘.

### 2.2.2 Stand der Forschung

Einer der ersten Interest-Operatoren ist bereits 1977 von MORAVEC [MOR77] entwickelt worden. Seitdem gibt es eine Vielzahl an Veröffentlichungen zum Thema Interest-Operatoren und Punkt-Detektoren. Eine umfassende Übersicht der aktuellen Methoden zur Extraktion von Punkt-Merkmalen ist z.B. in KÖTHE [KÖT03] sowie SCHMID u.a. [SCH00] zu finden. Die existierenden Verfahren können dabei prinzipiell in drei Kategorien eingeteilt werden:

- **Intensitätsbasiert:** Die Berechnung der Kenngröße für das Vorhandensein einer markanten Struktur erfolgt direkt aus den Intensitätswerten.
- **Konturbasiert:** Diese Methoden extrahieren zunächst eine Kontur und suchen anschließend nach Stellen mit maximaler Krümmung oder führen eine polygonale Approximation des Konturverlaufs durch und detektieren anschließend die Schnittpunkte.
- **Modellbasiert:** Durch Anpassen parametrischer Intensitätsmodelle an das Bildmuster kann für die Lokalisierung eine Subpixel-Genauigkeit erreicht werden.

Bei den konturbasierten Methoden sind Punkte im Bereich von Kreuzungen kritisch, da durch die Kantenextraktion an Stellen, wo sich drei oder mehr Kanten treffen, Unterbrechungen oder falsche Verbindungen entstehen können [CAN86]. Die modellbasierten Varianten sind je nach der verwendeten Wissensbasis auf sehr spezielle Strukturen (z.B. L-Ecken) beschränkt. In der Praxis sind für den universellen Einsatz die intensitätsbasierten Methoden am weitesten verbreitet.

Von SCHMID u.a. [SCH00] wurde ein praxisnaher Vergleich aktueller Interest-Operatoren unter Verwendung der Original-Implementierungen durchgeführt. Bei der aufwändigen quantitativen Evaluierung der Operatoren von FÖRSTNER [FÖR94], COTTIER, HEITGER, HORAUD sowie HARRIS und STEPHENS [HAR88] wurden im Wesentlichen die beiden folgenden Kriterien untersucht:

- **Wiederholbarkeit:** Neben der Robustheit gegenüber Bildrauschen wurde auch die Invarianz der Punktextraktion bei unterschiedlichen Rotationen, Skalierungen, Blickwinkeln und Beleuchtungssituationen getestet.
- **Informationsgehalt:** Die charakteristischen Eigenschaften der Interest-Werte wurden durch Berechnung der Entropie statistisch bewertet.

Zusammenfassend bleibt die Auswahl eines optimalen Verfahrens aber immer noch schwierig, da die vorgestellten Resultate erheblich von der jeweiligen Implementierung abhängig sind. Um die methodischen Vorteile besser herausarbeiten zu können, wurden die zwei führenden Interest-Operatoren auf der Grundlage des PLESSEY Punkt-Detektors und des FÖRSTNER-Operators unter vergleichbaren Bedingungen reimplementiert.

### 2.2.3 PLESSEY Punkt-Detektor

HARRIS und STEPHENS beschrieben in [HAR88] eine Verbesserung des klassischen MORAVEC-Operators [MOR77]. Sie lösten das Problem der diskreten Verschiebungen und Richtungen mit Hilfe der *Autokorrelations-Funktion* und steigerten dadurch die Genauigkeit bei der Lokalisierung.

Die Idee der Autokorrelation besteht in einem statistischen Ähnlichkeitsvergleich (*Korrelation*) von einem gegenüber dem ursprünglichen Bild leicht verschobenen Bildausschnitt. Dieser enthält genau dann ein signifikantes Punktmerkmal, wenn die Ähnlichkeit für jede Verschiebung in der Umgebung abnimmt.

Die Autokorrelations-Matrix  $\mathbf{A}$  wird durch Summation der ersten Ableitung einer Bildfunktion  $f$  in dem Gebiet  $\Omega$  um einen Punkt berechnet

$$\mathbf{A}(x, y) = \begin{bmatrix} \sum_{(i,j) \in \Omega} f_x(i, j)^2 & \sum_{(i,j) \in \Omega} f_x(i, j) \cdot f_y(i, j) \\ \sum_{(i,j) \in \Omega} f_x(i, j) \cdot f_y(i, j) & \sum_{(i,j) \in \Omega} f_y(i, j)^2 \end{bmatrix}, \quad (2.9)$$

wobei die partiellen Ableitungen  $f_x$  und  $f_y$  im Gegensatz zur Original-Beschreibung mit den kontinuierlichen Differential-Operatoren aus Abschnitt 2.1.4 bestimmt werden. In diesem Fall kann die Größe des Einflussbereichs durch die verwendete Standardabweichung  $\sigma$  automatisch ermittelt werden. Die Matrix  $\mathbf{A}$  enthält schließlich die Nachbarschaftsstruktur an der Stelle  $(x, y)$  und zeichnet sich durch folgende Eigenschaften aus:

- **Rang 2:** Bei vollem Rang wird ein markanter *Punkt* gekennzeichnet.
- **Rang 1:** Eine singuläre Matrix lässt auf eine gerade *Kante* schließen.
- **Rang 0:** Die Matrix definiert eine homogene unstrukturierte *Fläche*.

Bei dem PLESSEY Punkt-Detektor wird die *Punktstärke*  $w$  aus der Autokorrelations-Matrix  $\mathbf{A}$  mit der folgenden Kenngröße ermittelt:

$$w = \det(\mathbf{A}) - k \cdot \text{spur}(\mathbf{A})^2 \quad (2.10)$$

Um eine günstige Trennung der Kanten von den Punkten zu erhalten, wird für den freien Parameter  $k$  ein Wert von 0,04 gewählt. Auf diese Weise ergeben sich bei Punkten positive und bei geraden Kanten negative Werte. Die Position des Interest-Punktes wird schließlich durch eine lokale Nicht-Maxima-Unterdrückung ermittelt.

## 2.2.4 FÖRSTNER-Operator

FÖRSTNER [FÖR94, FÖR87] sucht markante Punkte ebenfalls durch Anwenden der Autokorrelations-Matrix  $\mathbf{A}$ . Die im Original verwendeten diskreten Filtermasken sind für die Implementierung durch kontinuierliche Filterfunktionen ersetzt worden (vgl. Abschnitt 2.1.4).

Im Gegensatz zu HARRIS und STEPHENS betrachtet FÖRSTNER als Maß für den Interest-Wert die beiden Eigenwerte  $\lambda_1$  und  $\lambda_2$  von  $\mathbf{A}$  als Halbachsen einer *Fehlerellipse*. Durch die Berechnung ihrer *Größe*

$$w = \frac{1}{\lambda_1 + \lambda_2} = \frac{\det(\mathbf{A})}{\text{spur}(\mathbf{A})}, \quad w > 0 \quad (2.11)$$

und der *Rundheit*

$$q = 1 - \left( \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \right)^2 = \frac{4 \cdot \det(\mathbf{A})}{\text{spur}(\mathbf{A})^2}, \quad 0 \leq q \leq 1 \quad (2.12)$$

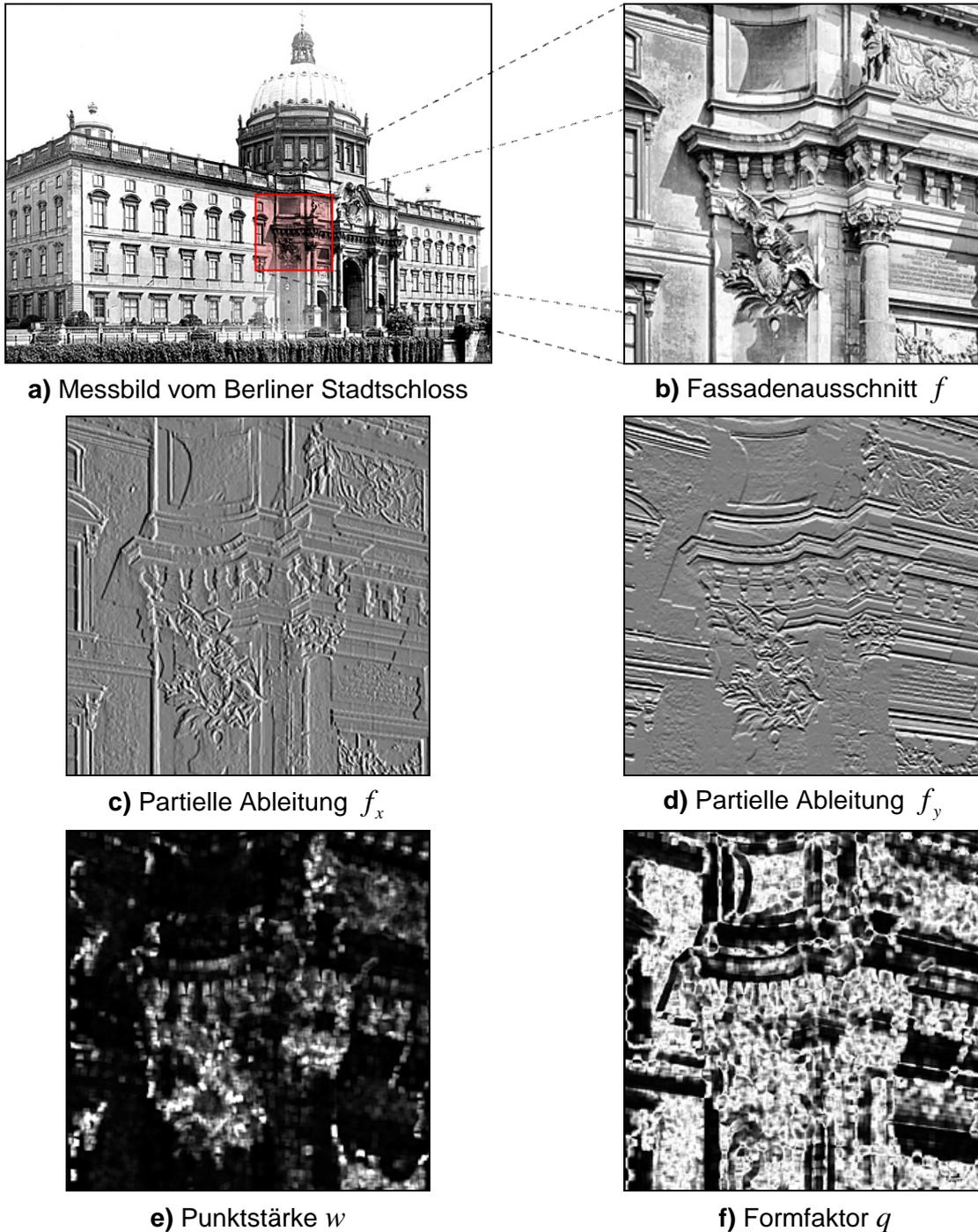
können schließlich charakteristische Merkmale abgeleitet werden:

- Kleine kreisförmige Ellipsen definieren einen markanten *Punkt*.
- Langgestreckte Fehlerellipsen lassen auf eine gerade *Kante* schließen.
- Große Ellipsen kennzeichnen eine homogene unstrukturierte *Fläche*.

Ein Interest-Punkt liegt genau dann vor, wenn die vorgegebenen Schwellwerte  $w_{\min}$  und  $q_{\min}$  überschritten werden. Geeignete Parameter hierfür bewegen sich in dem Bereich

$$w_{\min} = (0,5 \dots 1,5) \cdot \bar{w} \quad \text{und} \quad q_{\min} = 0,5 \dots 0,75, \quad (2.13)$$

wobei  $\bar{w}$  den Mittelwert von  $w$  über das gesamte Bild bezeichnet. Zur Veranschaulichung sind die beim FÖRSTNER-Operator produzierten Zwischenergebnisse in Abb. 2.5 und das Resultat der Merkmalsextraktion in Abb. 2.15 dargestellt.



**Abb. 2.5:** Zwischenergebnisse des FÖRSTNER-Operators am Beispiel eines Fassadenausschnitts des Berliner Stadtschlusses

### 2.2.5 Vergleich der Methoden

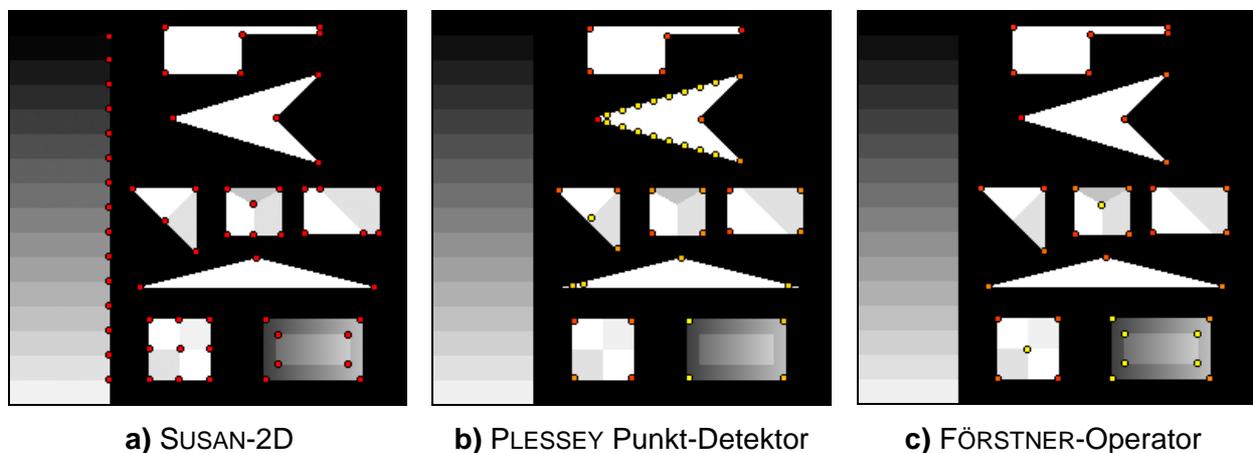
Nachfolgend sollen die Implementierungen des PLESSEY Punkt-Detektors und des FÖRSTNER-Operators gegenübergestellt werden. Für einen unabhängigen Vergleich wird das aufgrund des frei verfügbaren Quellcodes beliebte Verfahren SUSAN-2D (*Smallest Univalve Segment Assimilation Nucleus*) von SMITH und BRADY [SMI97] ohne Modifikationen herangezogen.

Im Wesentlichen vergleicht der Algorithmus die Helligkeit von jedem Pixel innerhalb einer kreisförmigen Maske mit dem mittleren Pixel, um ein Gebiet zu bestimmen, das eine ähnliche Helligkeit wie die Mitte besitzt. In diesem Gebiet werden durch Bestimmen der Größe, des Zentroiden und des zweiten Moments Punktmerkmale abgeleitet.

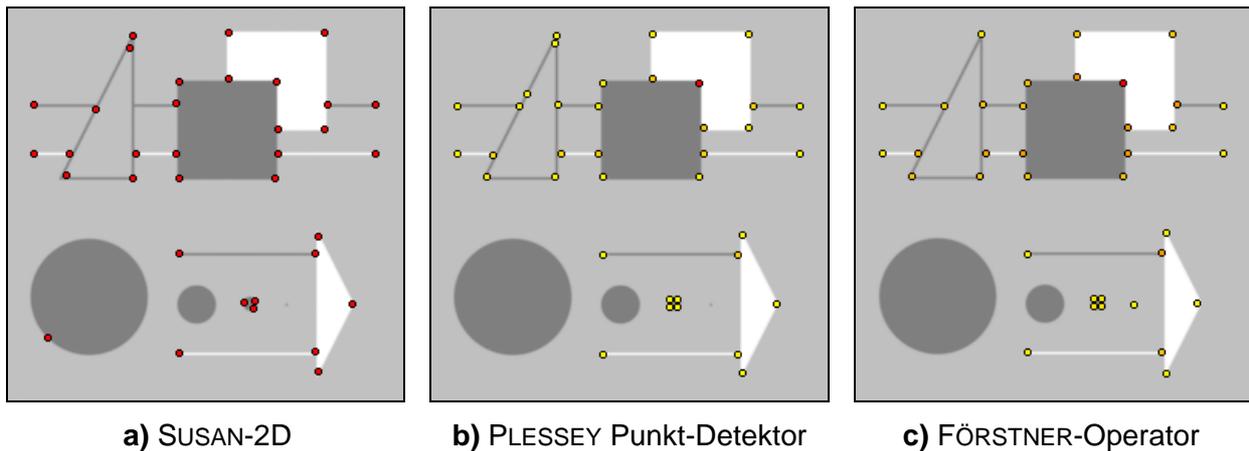
Zunächst soll die Leistungsfähigkeit der drei Testkandidaten bei der Anwendung auf standardisierte Referenzbilder verglichen werden. Die Ergebnisse der Merkmalsextraktion für die synthetischen Testbilder sind in Abb. 2.6 und Abb. 2.7 dargestellt. Dabei entspricht die Färbung der Interest-Punkte bei den eigenen Implementierungen der berechneten Punktstärke.

Die Vorteile des SUSAN-2D liegen in der schnellen Berechnung und den hervorragenden Ergebnissen bei ungestörten und nicht geglätteten Bildern, insbesondere dem eigenen Testbild (siehe Abb. 2.6). Bei dem PLESSEY Punkt-Detektor fällt auf, dass das Verfahren auf die scharfen Rasterstrukturen des Bildes sehr empfindlich reagiert. Nachteilig ist auch, dass für die verwendete Definition der Punktstärke kein Schwellwert gefunden werden konnte, der die wichtigen von den uninteressanten Punkten trennt.

Die beim FÖRSTNER-Operator betrachtete Fehlerellipse als Maß für den Interest-Wert zeigt hier deutlich bessere Resultate. Durch die Anforderung einer annähernd runden Ellipse mit  $q_{\min} = 0,5$  werden allerdings Punkte an geraden Kanten unterdrückt, so dass insgesamt weniger Merkmale detektiert werden als z.B. beim SUSAN-2D.



**Abb. 2.6:** Ergebnisse der Interest-Operatoren mit dem Referenzbild von SMITH und BRADY



**Abb. 2.7:** Vergleich der Interest-Operatoren bei dem Referenzbild von ROSENTHALER

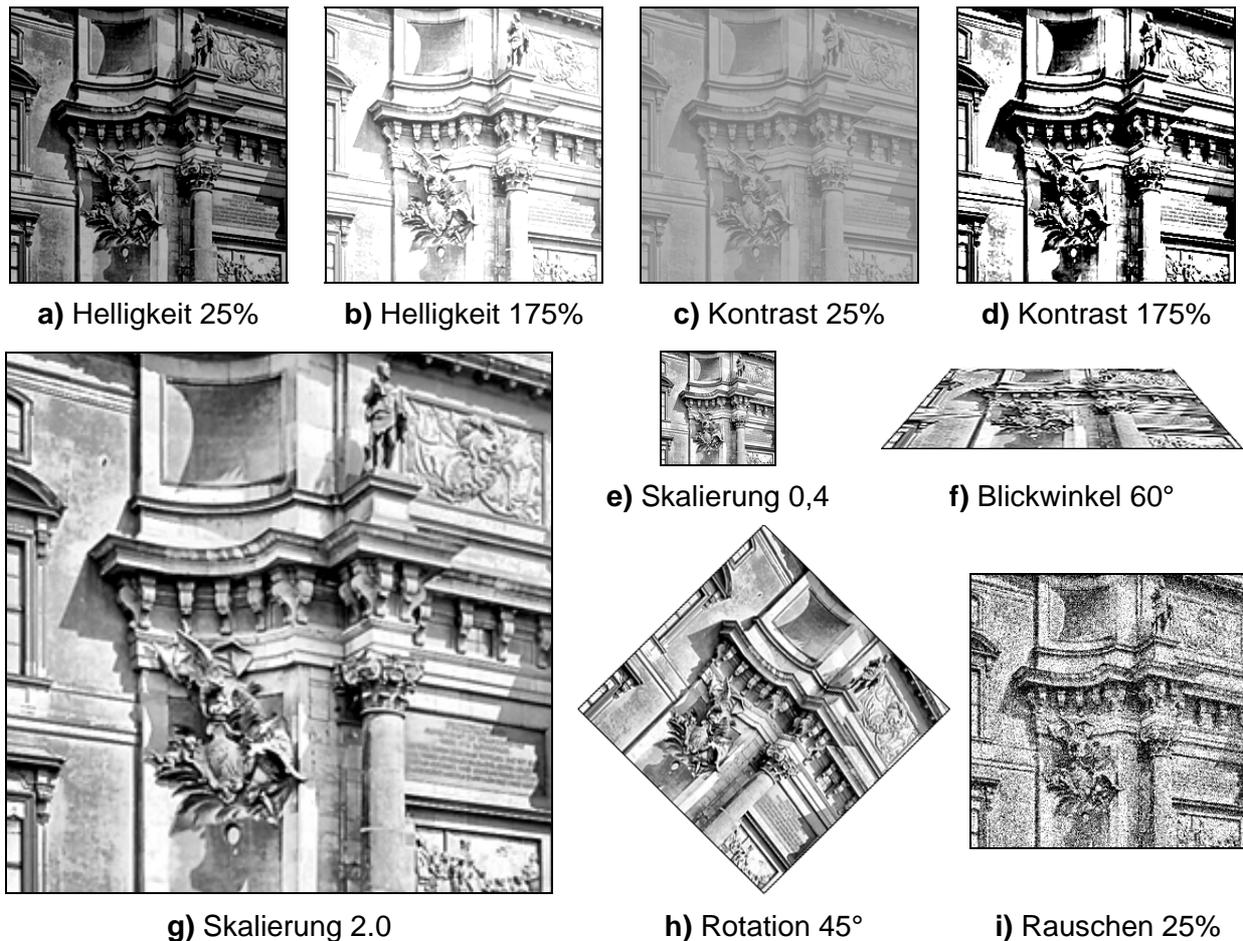
Das synthetische Testbild von ROSENTHALER u.a. [ROS92] simuliert etwas realistischere Bedingungen, da es weniger Kontrast enthält und ein wenig geglättet ist (vgl. Abb. 2.7). Wie erwartet zeigen die beiden Verfahren mit der robusten Gradientenberechnung bessere Ergebnisse im Vergleich zum SUSAN-2D.

Allerdings ist es bei den Experimenten nur mit dem FÖRSTNER-Operator gelungen, alle potentiellen Interest-Punkte ohne Fehler zu finden. Diese Einschätzung soll durch einen quantitativen Vergleich verifiziert und gleichzeitig die Stabilität der Merkmalsextraktion abgeschätzt werden.

Dazu wurde der Fassadenausschnitt mit unterschiedlichen Transformationen verändert (vgl. Abb. 2.8):

- **Helligkeit und Kontrast:** Die radiometrischen Eigenschaften wurden gleichmäßig für das gesamte Bild jeweils zwischen 25 und 175 Prozent variiert.
- **Rotation:** Die Drehungen in der Bildebene sind aufgrund der Symmetrie nur in Schritten zwischen 0 und 90 Grad durchgeführt worden.
- **Blickwinkel:** Verschiedene Ansichten wurden durch die perspektivische Abbildung eines um die horizontale Achse rotierten Bildes zwischen 0 und 60 Grad simuliert.
- **Skalierung:** Die Bildgröße ist mit einem Faktor zwischen 0,4 und 2,0 verändert worden.
- **Rauschen:** Die Intensitätswerte des Bildes wurden mit einer GAUSSschen Normalverteilung bis zu 25 Prozent gestört.

Als Kriterium für die Evaluierung wurde die Wiederholungsrate gewählt. Ausgehend von den bekannten Sollstellen der Interest-Punkte in den Testbildern ist die Anzahl der korrekten Punkte in den transformierten Bildern visuell ausgezählt worden.

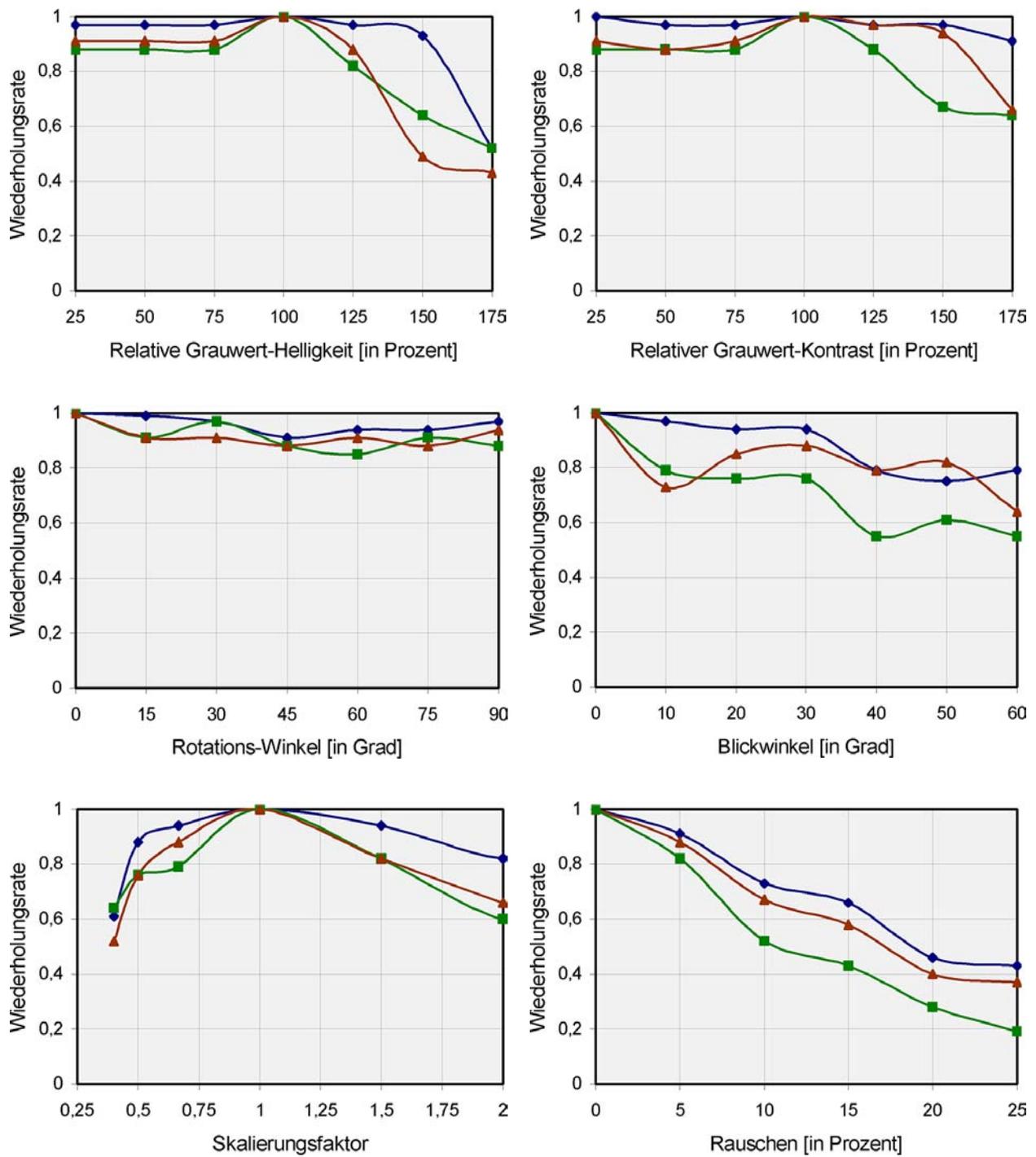


**Abb. 2.8:** Testbilder für den quantitativen Vergleich der Interest-Operatoren

Die detaillierten Ergebnisse dieses quantitativen Vergleichs sind in Abb. 2.9 dargestellt. Die Untersuchungen haben gezeigt, dass z.B. bei den radiometrischen Änderungen der FÖRSTNER-Operator eine doppelt so starke Aufhellung des Bildes verkraftet wie der PLESSEY Punkt-Detektor. Des Weiteren führt eine Erhöhung des Kontrastes beim SUSAN-2D sofort zu Ausfällen, während das Ergebnis beim FÖRSTNER-Operator nahezu konstant bleibt.

Alle drei Testkandidaten zeigen bei Drehungen in der Bildebene erfreulich stabile Resultate. Allerdings können bei der perspektivischen Abbildung oder zunehmenden Größenänderung zwischen den Verfahren erhebliche Differenzen beobachtet werden. Auch wenn die Stabilität aller Verfahren bei den durch Rauschen gestörten Bildern signifikant nachlässt, ist der FÖRSTNER-Operator dem SUSAN-2D deutlich überlegen.

Zusammenfassend zeigt die Implementierung des FÖRSTNER-Operators bei den durchgeführten Experimenten die besten Ergebnisse.



**Abb. 2.9:** Vergleich der Wiederholungsraten bei den transformierten Testbildern für den SUSAN-2D (■), den PLESSEY Punkt-Detektor (▲) und den FÖRSTNER-Operator (◆)

## 2.3 Verbesserungen und Erweiterungen

Nachdem sich der FÖRSTNER-Operator unter den Anforderungen der Bildzuordnung als leistungsfähigstes Verfahren durchgesetzt hat, sollen im Folgenden noch diverse Verbesserungen vorgeschlagen werden. Um die Genauigkeit bei der Lokalisierung zu erhöhen, werden zunächst zwei verschiedene Techniken zur Subpixel-Bestimmung vorgestellt und bewertet (Abschnitt 2.3.1).

Im Anschluss wird ein adaptives Schwellwertverfahren vorgeschlagen, damit Punkthäufungen vermieden und eine gleichmäßige Verteilung im Bild sichergestellt werden kann (Abschnitt 2.3.2). Abschließend wird die Erweiterung der Autokorrelation für den Einsatz von Farbbildern beschrieben (Abschnitt 2.3.3).

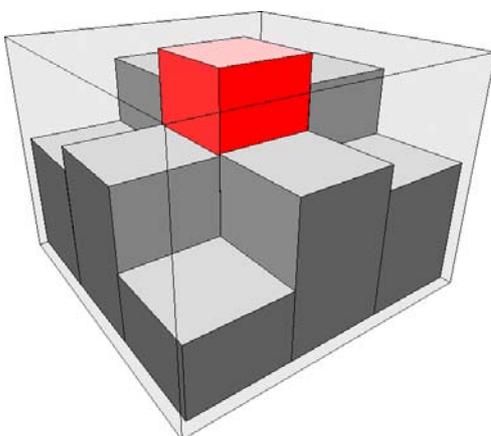
### 2.3.1 Subpixel-Bestimmung durch Paraboloid-Anpassung

Da die Genauigkeit bei der Bestimmung der Bildkoordinaten einen entscheidenden Einfluss auf die spätere Rekonstruktion hat, sollte die Lokalisierung möglichst präzise erfolgen. Des Weiteren kann bei der merkmalsbasierten Korrespondenz-Analyse das Deformations-Modell mit einer exakten Position weitaus stabiler berechnet werden (vgl. Abschnitt 3.3).

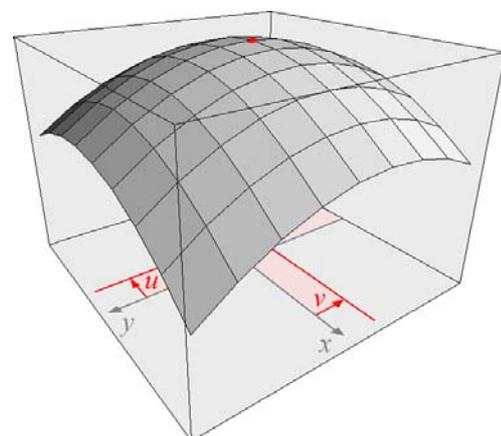
Deshalb soll für punktförmige Merkmale die Subpixel-Position mit Hilfe eines *Paraboloid-Modells* bestimmt werden. Ausgehend von der ganzzahligen Position  $(x, y)$  eines Punktes mit maximaler Punktstärke  $w$  wird seine unmittelbare Nachbarschaft betrachtet (siehe Abb. 2.10) und die neun Werte auf 1 normiert. Zur Modellierung kann ein Paraboloid mit der biquadratischen Funktion

$$w(x, y) = ax^2 + by^2 + cxy + dx + ey + f \quad (2.14)$$

konstruiert werden, dessen Oberfläche auf die normierte Punktstärke angepasst wird.



a) Lokales Maximum der Punktstärke



b) Angepasste Oberfläche des Paraboloids

**Abb. 2.10:** Räumliche Darstellung einer Paraboloid-Anpassung an die Punktstärke

Anschaulich entspricht diese Oberfläche zweiter Ordnung einer Parabel, die um ihre Symmetrieachse rotiert wird und dabei so in der Breite variiert werden kann, dass sie ein Paraboloid mit einer elliptischen Basis bildet.

Zunächst soll ein *numerischer Ansatz* zur Anpassung des Paraboloids beschrieben werden. Um die Koeffizienten des Paraboloids  $\mathbf{x} = (a, \dots, f)^T$  zu bestimmen, kann ein lineares Gleichungssystem der Form

$$\mathbf{A} \mathbf{x} = \mathbf{b} \quad (2.15)$$

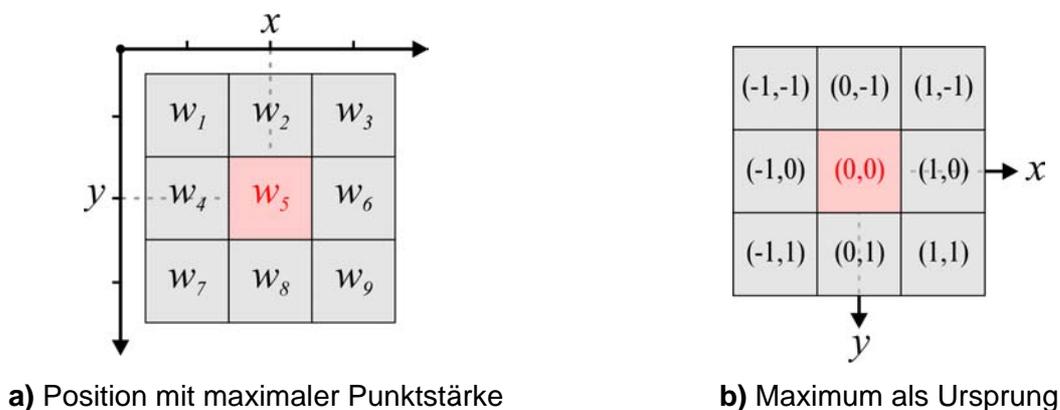
aufgestellt werden, wobei

$$\mathbf{A} = \begin{bmatrix} x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 & 1 \\ x_2^2 & y_2^2 & x_2 y_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_9^2 & y_9^2 & x_9 y_9 & x_9 & y_9 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{b} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_9 \end{bmatrix} \quad (2.16)$$

bezeichnen. Wenn ein Punkt mit maximaler Punktstärke an der Stelle  $(x, y)$  als Ursprung für ein lokales Koordinatensystem definiert wird (siehe Abb. 2.11), dann vereinfachen sich die Parameter von  $\mathbf{A}$  zu einer konstanten Design-Matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \quad (2.17)$$

Die Verwendung des maximalen Punktes mit seinen acht Nachbarn führt bei nur sechs Unbekannten in dem Koeffizienten-Vektor  $\mathbf{x}$  zu einem überbestimmten linearen Gleichungssystem, das numerisch z.B. mit Hilfe der *Singulärwert-Zerlegung* (Singular Value Decomposition, SVD) gelöst werden kann (vgl. Anhang A.2.1).



**Abb. 2.11:** Definition des lokalen Koordinatensystems

Eine weitere Möglichkeit zur Bestimmung der Koeffizienten des Paraboloids besteht in der Anwendung von *lokalen Operatoren* auf die Punktstärke  $w$ . Die Betrachtung der partiellen Ableitungen der Paraboloid-Funktion zeigt zunächst folgenden Zusammenhang:

$$\begin{aligned} w_{xx} = \frac{\partial w}{\partial x^2} = 2a, \quad w_{yy} = \frac{\partial w}{\partial y^2} = 2b, \quad w_{xy} = \frac{\partial w}{\partial x \partial y} = c, \\ w_x = \frac{\partial w}{\partial x} = 2ax + cy + d \quad \text{und} \quad w_y = \frac{\partial w}{\partial y} = 2by + cx + e \end{aligned} \quad (2.18)$$

Da jeder Punkt in seinem lokalen Koordinatensystem den Ursprung bildet, vereinfachen sich die gesuchten Koeffizienten des Paraboloids zu

$$a = \frac{w_{xx}}{2}, \quad b = \frac{w_{yy}}{2}, \quad c = w_{xy}, \quad d = w_x \quad \text{und} \quad e = w_y. \quad (2.19)$$

Wie bereits in Abschnitt 2.1.4 beschrieben, können die partiellen Ableitungen erster und zweiter Ordnung relativ einfach durch Faltung an der Stelle des lokalen Maximums  $(x, y)$  approximiert werden.

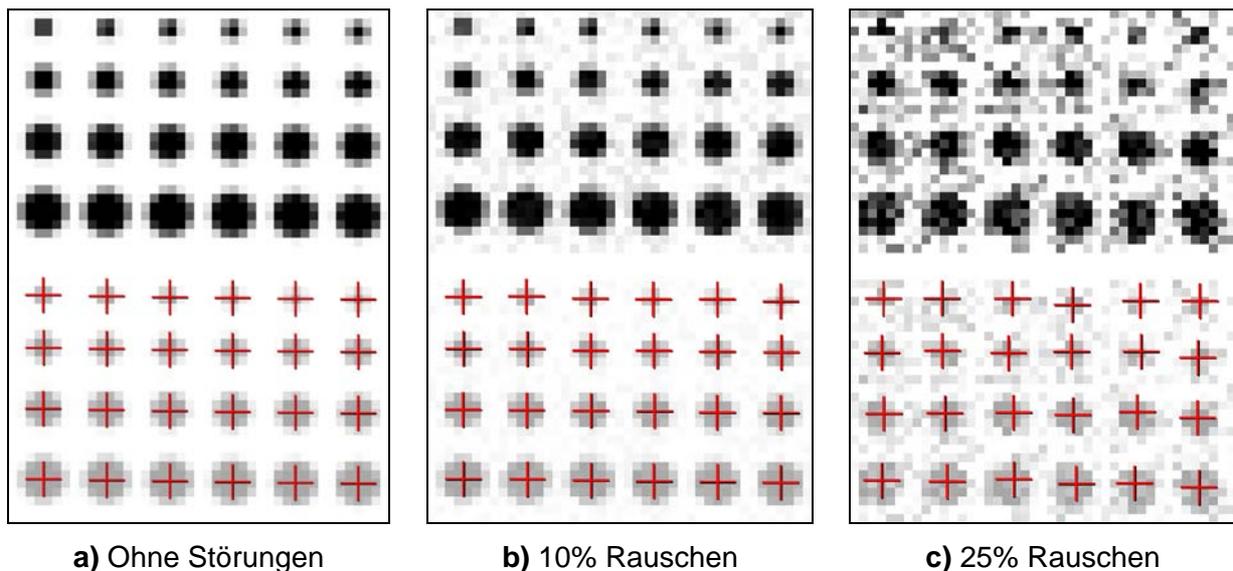
Nachdem die Koeffizienten des Paraboloids mit einem der beiden Verfahren berechnet wurden, kann die Position für das Maximum gesucht werden (siehe Abb. 2.10). Hierfür müssen die ersten partiellen Ableitungen  $w_x$  und  $w_y$  aus der Gleichung (2.18) gleich 0 gesetzt werden. Die Subpixel-Verschiebung in  $x$ - und  $y$ -Richtung kann durch Einsetzen direkt in geschlossener Form gelöst werden. Nach dem Umformen gilt

$$u = \frac{(2bd - ce)}{(c^2 - 4ab)} \quad \text{und} \quad v = \frac{(2ae - cd)}{(c^2 - 4ab)}. \quad (2.20)$$

Abschließend kann durch Addition der Subpixel-Verschiebung zu der ganzzahligen Position  $(x + u, y + v)$  die Lagebestimmung für punktförmige Bildmerkmale verbessert werden.

Um die Genauigkeit der Subpixel-Bestimmung bewerten zu können, wurden synthetische Testmuster erzeugt (vgl. Abb. 2.12). Während in der oberen Bildhälfte Punkte zwischen zwei und fünf Pixeln an verschiedenen Stellen dargestellt sind, markieren die Kreuze in der unteren Hälfte die berechneten Subpixel-Positionen. Durch Hinzufügen von Rauschen einer GAUSSschen Normalverteilung mit 10 bzw. 25 Prozent sollen realistische Bedingungen angenähert und die Robustheit demonstriert werden.

Die Tab. 2.1 enthält die quantitativen Ergebnisse der beiden Methoden zur Subpixel-Bestimmung. Sie zeigt, dass bei dem verwendeten Testbild unter idealen Bedingungen die Position mit einer durchschnittlichen Genauigkeit von ungefähr 1/36 Pixel bestimmt werden kann und eine maximale Abweichung von 1/18 Pixel nicht überschritten wird.



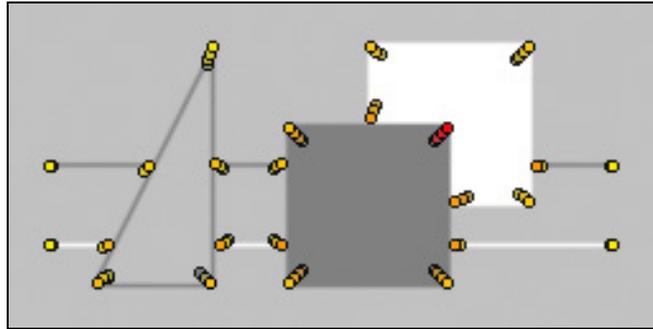
**Abb. 2.12:** Robustheit der Subpixel-Bestimmung bei unterschiedlich starken Störungen

Bei stark gestörten Bildern mit einem Rauschanteil von 25 Prozent kann sich die mittlere Auflösung auf  $1/6$  Pixel und die Zuverlässigkeit sogar auf ein halbes Pixel reduzieren. Beide Verfahren zeigen vergleichbare Resultate, wobei die numerische Subpixel-Bestimmung den lokalen Operatoren leicht überlegen ist.

Die Verwendung eines Fensters endlicher Größe zur Autokorrelation führt jedoch bei L- oder T-Ecken zu einem systematischen Lagefehler (vgl. Abb. 2.13). Die Verschiebung in Bezug auf die bei der GAUSS-Funktion verwendete Standardabweichung  $\sigma$  ist z.B. bei WANG und BRADY [WAN94] beschrieben. FÖRSTNER und GÜLCH [FÖR87] sowie DERICHE und GIRAUDON [DER93] schlagen als Lösung ein mehrstufiges Verfahren vor, wobei durch Schneiden von Kanten versucht wird, die Genauigkeit der Lagebestimmung zu erhöhen.

Rauschen [in Prozent]	Mittlerer Fehler [in Pixel]				Maximaler Fehler [in Pixel]				Auflösung [in Pixel]	Zuverlässigkeit [in Pixel]
	2	3	4	5	2	3	4	5		
<b>Numerische Berechnung</b>										
<b>0</b>	0,029	0,030	0,024	0,023	0,052	0,045	0,045	0,059	<b>1/37</b>	<b>1/19</b>
<b>10</b>	0,058	0,038	0,043	0,037	0,116	0,104	0,115	0,175	<b>1/22</b>	<b>1/7</b>
<b>25</b>	0,147	0,148	0,116	0,135	0,462	0,668	0,424	0,481	<b>1/7</b>	<b>1/2</b>
<b>Lokale Operatoren</b>										
<b>0</b>	0,032	0,032	0,025	0,024	0,057	0,051	0,046	0,060	<b>1/35</b>	<b>1/18</b>
<b>10</b>	0,059	0,045	0,043	0,038	0,122	0,112	0,111	0,185	<b>1/21</b>	<b>1/7</b>
<b>25</b>	0,163	0,156	0,138	0,170	0,468	0,451	0,422	0,456	<b>1/6</b>	<b>1/2</b>

**Tab. 2.1:** Lokalisierungs-Fehler der beiden Subpixel-Verfahren in Abhängigkeit von Punktgröße und Rauschanteil



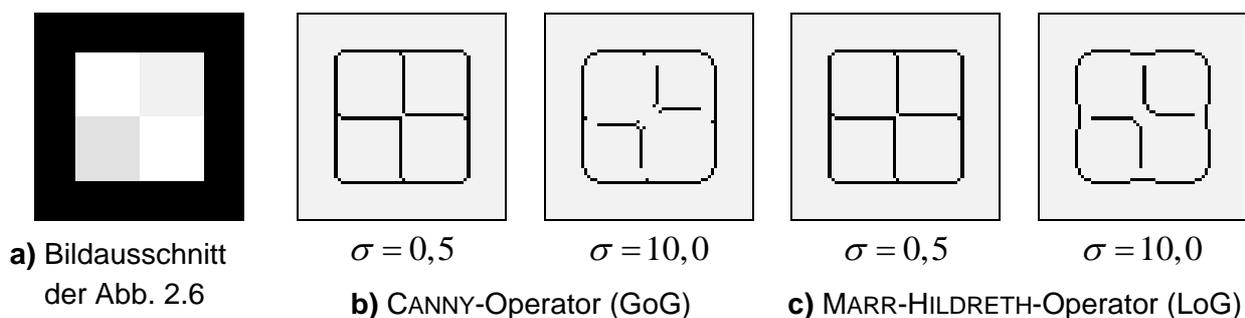
**Abb. 2.13:** Systematische Lagefehler bei unterschiedlich starker Glättung

Klassische Kanten-Operatoren, die das Maximum der ersten Ableitung suchen (*Gradient of GAUSSian*, GoG), sind für eine Korrektur allerdings weniger geeignet, da sie die gleichen Verschiebungen aufweisen. LoG-Operatoren (*Laplacian of GAUSSian*), die Nulldurchgänge der zweiten Ableitung suchen, erschweren das Schneiden von Kanten durch starkes Abrunden der interessanten Ecken. Dieses Verhalten ist in Abb. 2.14 am Beispiel der Kanten-Operatoren von CANNY [CAN86] bzw. MARR und HILDRETH [MAR80] dargestellt.

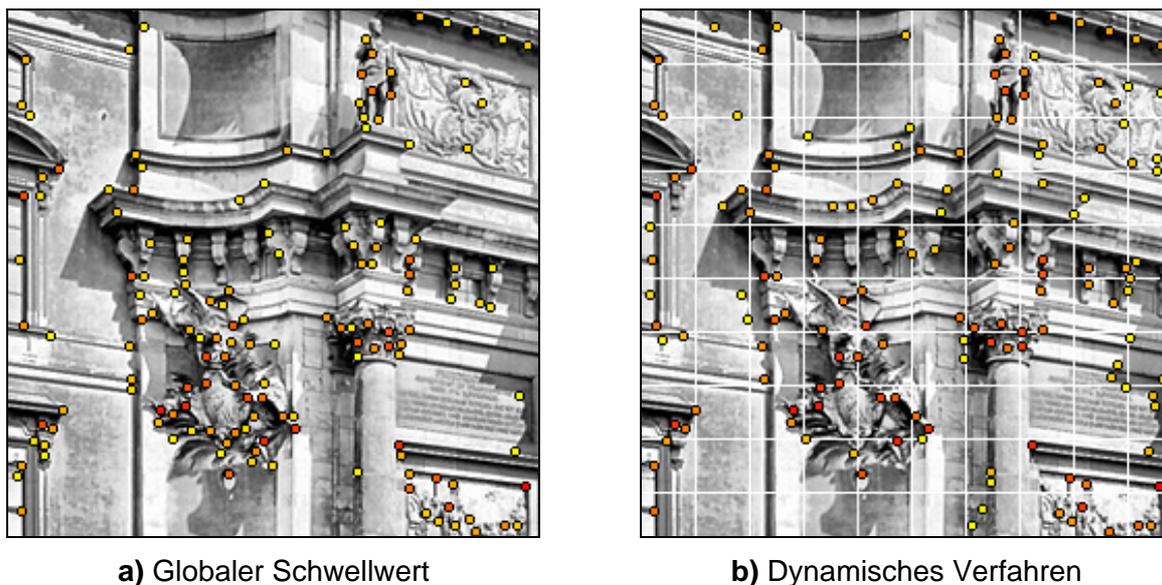
### 2.3.2 Gleichmäßige Punktverteilung mit adaptivem Schwellwert

Aufgrund der Anforderungen für die Bildzuordnung (vgl. Abschnitt 2.2.1) werden nicht alle extrahierten Punktmuster benötigt, sondern es sollten nur die wichtigsten Interest-Punkte ausgewählt werden. Bei der Verwendung eines globalen Schwellwertes für das gesamte Bild kann es allerdings in strukturierten Bereichen zu ungünstigen Punktanhäufungen kommen, während homogene Bereiche vernachlässigt werden. Diese ungleichmäßige Verteilung der Interest-Punkte wirkt sich nachteilig auf das in Kapitel 4 beschriebene Orientierungs-Verfahren aus.

Für die Optimierung der Punktverteilung kann ein *adaptives Schwellwert-Verfahren* eingesetzt werden. Um die lokale Punktdichte zu beeinflussen, wird zunächst das gesamte Bild in möglichst gleich große Segmente eingeteilt.



**Abb. 2.14:** Problem der Lagebestimmung durch Schneiden von Kanten



**Abb. 2.15:** Optimierung der Punktverteilung durch einen adaptiven Schwellwert

Gemäß der auszuwählenden Punktzahl erhält dann jedes Segment anteilig ein Kontingent, das für die stärksten Punkte in diesem Bereich verwendet wird. Verbleibt bei der Aufteilung ein Restkontingent oder enthalten Segmente weniger Punkte als im Kontingent vorgesehen, werden diese anschließend an die global stärksten Punkte vergeben. Abb. 2.15 zeigt die Optimierung der Punktverteilung mit dem adaptiven Schwellwert-Verfahren am Beispiel des Fassadenausschnitts aus Abschnitt 2.2.4. Dabei sind von ursprünglich ca. 1.000 extrahierten Punktmustern die wichtigsten 150 Kandidaten in 100 Segmenten ausgewählt worden.

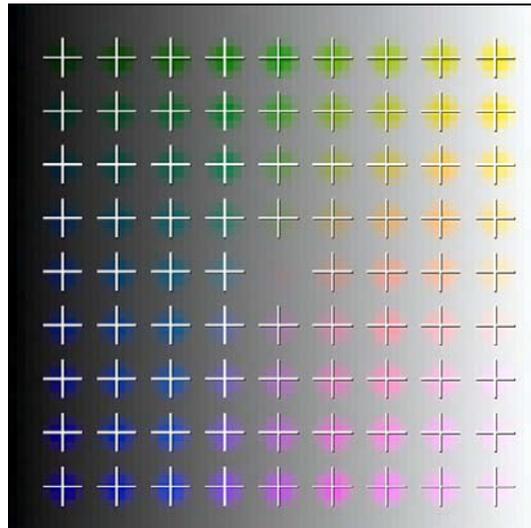
### 2.3.3 Erweiterung der Autokorrelation für Farbbilder

Mit einem Großteil der aktuellen Aufnahmesysteme können farbige Bilder erzeugt werden (vgl. Abschnitt 1.4.2). Durch eine Erweiterung des Interest-Operators für Farbbilder sollen Punktmerkmale, die sich zwar nicht in der Helligkeit von dem Hintergrund absetzen, aber durch ihren Farbton in der Umgebung auffallen, ebenfalls erkannt werden.

MONTESINOS u.a. [MON98] erweiterten die Autokorrelations-Matrix  $\mathbf{A}$  für Farbbilder  $f = (r, g, b)^T$ , wobei zunächst für jeden Farbkanal die erste partielle Ableitung berechnet werden muss. Die Bestimmung der Autokorrelations-Matrix erfolgt dann analog zum Abschnitt 2.2.3 mit

$$\mathbf{A}_{RGB}(x, y) = \begin{bmatrix} \sum (r_x^2 + g_x^2 + b_x^2) & \sum (r_x r_y + g_x g_y + b_x b_y) \\ \sum (r_x r_y + g_x g_y + b_x b_y) & \sum (r_y^2 + g_y^2 + b_y^2) \end{bmatrix}, \quad (2.21)$$

wobei die Indices der Summen für das Gebiet  $\Omega$  der Einfachheit halber weggelassen worden sind.



**Abb. 2.16:** RGB-Testbild zur Extraktion farbiger Punktmerkmale

Um die Funktionsweise der Farberweiterung zu demonstrieren, wurde ein synthetisches Testmuster mit farbigen Punkten erstellt. Das RGB-Farbbild besitzt einen kontinuierlichen Helligkeitsverlauf und die Punkte unterscheiden sich nur in ihrem Farbton oder der Sättigung vom Hintergrund (siehe Abb. 2.16). Dabei kennzeichnet jede Kreuz-Markierung einen erfolgreich detektierten farbigen Interest-Punkt.

## 2.4 Experimentelle Ergebnisse

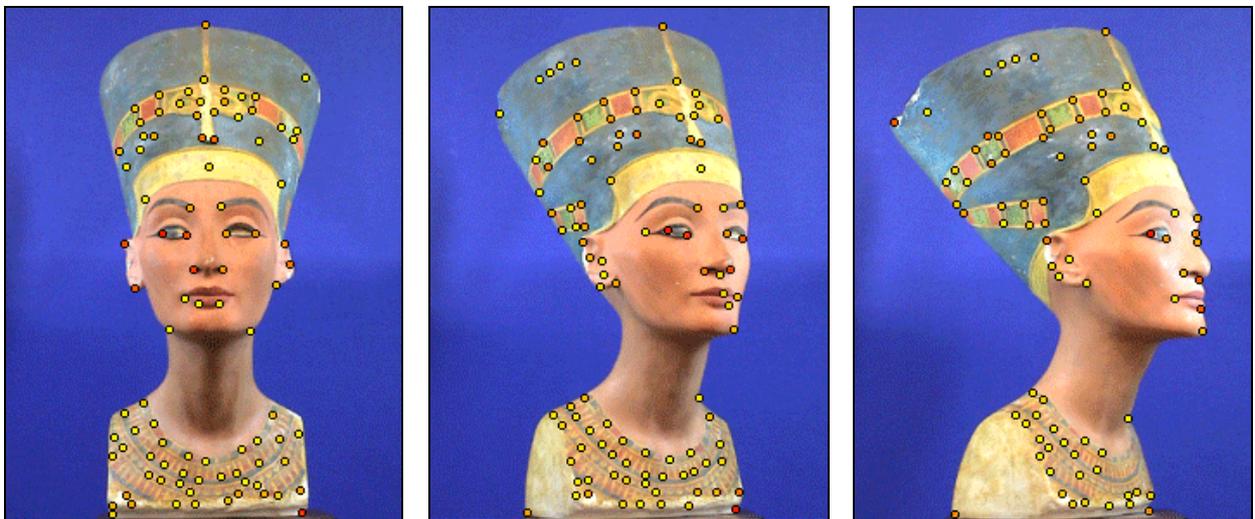
Nachdem die Wirkungsweise des implementierten Interest-Operators bereits an einigen synthetischen Testbildern veranschaulicht wurde, sollen nun Ergebnisse für reale Bildbeispiele vorgestellt werden. Um die Darstellung bei den hochauflösenden Messbildern der MEYDENBAUER-Kamera übersichtlich zu halten (siehe Abb. 2.17), sind die digitalisierten Bildausschnitte vor der Merkmalsextraktion mit dem Faktor  $2^5$  in einer Auflösungspyramide verkleinert worden (vgl. Abschnitt 3.5.2).

Bei den historischen Fassadenausschnitten basieren die Interest-Punkte ausschließlich auf Grauwerten. Ergebnisse für Farbbilder können am Beispiel der NOFRETETE demonstriert werden (siehe Abb. 2.18). Die Aufnahmen wurden mit einem DV-Camcorder in einer Auflösung von  $1.020 \times 1.360$  Pixeln erzeugt [KUZ02].

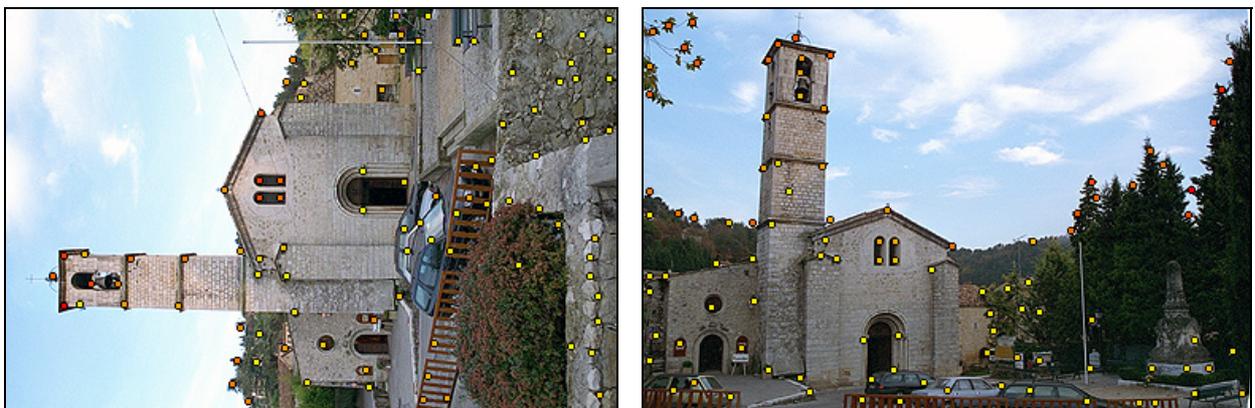
Schließlich sind aus der häufig verwendeten VALBONNE-Sequenz [SCH01] noch zwei um 90 Grad gedrehte Testbilder mit  $768 \times 512$  Pixeln ausgewählt worden, bei denen der gemeinsam sichtbare Objektbereich nur einen kleinen Anteil des Bildes ausmacht (vgl. Abb. 2.19).



**Abb. 2.17:** Automatisch extrahierte Punktmerkmale am Eingang zum Berliner Stadtschloss



**Abb. 2.18:** Ergebnisse des Interest-Operators bei einer Nachbildung der NOFRETETE-Büste aus dem Ägyptischen Museum in Berlin



**Abb. 2.19:** Extrahierte Interest-Punkte bei den Testbildern der Kirche in VALBONNE, Frankreich

## 2.5 Diskussion

Die in diesem Kapitel dargestellten Verfahren zur digitalen Bildverarbeitung und zur Merkmalsextraktion mit Hilfe von Interest-Operatoren liefern stabile Punktmuster, die sich als Kandidaten für eine Bildzuordnung eignen. Die Implementierung der robusten und gut parametrisierbaren Gradienten-Berechnung mit den kontinuierlichen Differential-Operatoren liefert in der Praxis weitaus bessere Ergebnisse als die Anwendung von diskreten Filtermasken oder Grauwert-Differenzen. Der methodische Vergleich und die quantitative Analyse führender Interest-Operatoren haben gezeigt, dass der implementierte FÖRSTNER-Operator unter den Anforderungen der Bildzuordnung die besten Resultate erzielt.

Des Weiteren konnte nachgewiesen werden, dass bei isolierten Punkten durch eine Paraboloid-Anpassung die Lokalisierung mit Subpixel-Genauigkeit erreicht werden kann. Dabei haben die Experimente ergeben, dass die numerische Berechnung leichte Vorteile gegenüber der Verwendung von lokalen Operatoren besitzt. Um die Stabilität der nachfolgenden Orientierungs-Verfahren zu erhöhen, wurde ein adaptives Schwellwert-Verfahren vorgeschlagen, das die Punktverteilung im Bild optimiert. Damit die zusätzlichen Informationen von Farbbildern ausgewertet werden können, wurde schließlich auch eine entsprechende Erweiterung der Autokorrelations-Funktion für farbige Punktmuster beschrieben und getestet.

Allerdings sollte der systematische Lagefehler bei der Bestimmung von L- und T-Ecken reduziert werden, da aufgrund der perspektivischen Verzerrungen im Nahbereich die Verschiebungen in den Bildern unterschiedlich stark sind. Erste Versuche, den Lagefehler mit einem hierarchischen Ansatz in einer Bildpyramide zu beseitigen, zeigen zwar viel versprechende Resultate, die aber im Detail noch untersucht werden müssen. Abschließend wäre für die automatische Berechnung von Interest-Punkten eine lokale Statistik zur Schätzung des Steuerungs-Parameters sehr hilfreich.

*„Ich kann freilich nicht sagen, ob es  
besser wird, wenn es anders wird;  
aber soviel kann ich sagen,  
es muss anders werden,  
wenn es gut werden soll.“*

GEORG CHRISTOPH LICHTENBERG [1742-1799];  
dt. Physiker und Schriftsteller

### 3 Merkmalsbasierte Korrespondenz-Analyse

Nachdem in jedem Bild unabhängig voneinander Interest-Punkte bestimmt wurden, müssen diese nun einander zugeordnet werden. In der Photogrammetrie ist diese Aufgabe von entscheidender Bedeutung, da jedes automatisierte Triangulations-Messverfahren über eine leistungsfähige Korrespondenz-Analyse verfügen muss. Für die allgemeine Lösung des Korrespondenzproblems gibt es zahlreiche Ansätze. Eine umfassende Übersicht der existierenden Verfahren ist z.B. in FÖRSTNER [FÖR93] oder KOSCHAN [KOS91] zu finden.

Im Gegensatz zu den *flächenbasierten* Techniken erlauben die *merkmalsbasierten* Verfahren einen schnelleren Zuordnungsprozess bei einer geringeren Wahrscheinlichkeit von Fehlzuordnungen, da es weniger Korrespondenzkandidaten gibt. Sind Merkmalseigenschaften unterscheidbar, so wird man nur solche gleichen Typs und mit verträglichen Eigenschaften einander zuordnen. Bei einer *relationalen* Bildzuordnung können neben den extrahierten Bildmustern zusätzliche Beziehungen zwischen den Merkmalen (z.B. ‚ist parallel‘, ‚orthogonal zu‘, ‚liegt links‘ bzw. ‚oberhalb von‘) verwendet werden. Diese Relationen werden für jedes Bild einzeln aufgestellt und anschließend in die anderen Bilder transformiert. Das Einsatzgebiet für die relationale Bildzuordnung ist allerdings hauptsächlich im Bereich der modellgestützten Objekterkennung angesiedelt.

In der vorliegenden Arbeit besteht die Herausforderung darin, ohne Kenntnis der Abbildungsgeometrie eine initiale Menge von Punkt-Korrespondenzen zu finden. Dabei wird für jeden Interest-Punkt zunächst ein Referenzmuster definiert, das mit den Mustern potentieller Korrespondenzkandidaten verglichen werden kann (Abschnitt 3.1). Des Weiteren wird ein Kriterium benötigt, das die Ähnlichkeit der Muster möglichst robust und aussagekräftig beschreibt. In Abschnitt 3.2 werden hierzu verschiedene Korrelations-Methoden analysiert. Allerdings wird gerade im Nahbereich durch die perspektivische Abbildung ein Objekt je nach Aufnahme-richtung unterschiedlich verzerrt. Deshalb wird in Abschnitt 3.3 eine Kleinste-Quadrate-Korrelation vorgestellt, die verschiedene geometrische Deformationen kompensieren kann.

Das neu entwickelte Verfahren zur merkmalsbasierten Korrespondenz-Analyse basiert auf einer geeigneten Kombination dieser Techniken (Abschnitt 3.4). Die Mehrdeutigkeiten, die gerade bei kleinen Mustern und einer großen Anzahl von Korrespondenzkandidaten entstehen können, werden schließlich mit Hilfe einer hierarchischen Bildzuordnungs-Strategie gelöst (Abschnitt 3.5).

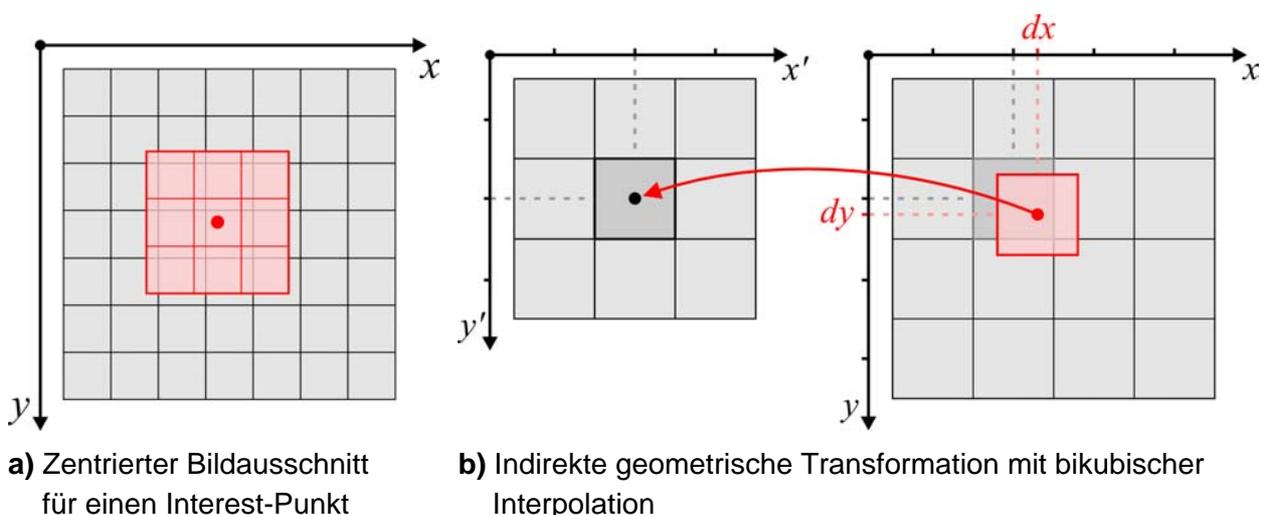
### 3.1 Definition der Vergleichsmuster

Für jeden Interest-Punkt muss ein Gebiet definiert werden, das für einen Ähnlichkeitsvergleich herangezogen werden kann. Aufgrund der im Nahbereich vorkommenden Deformationen sollte dieser Einflussbereich aber möglichst klein sein. Zur Vereinfachung wird für den zentrierten Bildausschnitt eine quadratische Form angenommen (siehe Abb. 3.1a).

Da die Position eines Interest-Punktes mit Subpixel-Genauigkeit berechnet wird, müssen die Intensitätswerte im Fenster entsprechend angepasst werden. Das implementierte Verfahren zur geometrischen Transformation wird neben dieser Subpixel-Verschiebung eines Bildfensters z.B. auch bei der Affin-Transformation in Abschnitt 3.3 und der Normalbild-Erstellung in Abschnitt 7.1 eingesetzt.

Durch Transformation der Bildkoordinaten kann ein neues digitales Bild mit veränderter geometrischer Lage erzeugt werden (*Entzerrung*). Bei einer *direkten* geometrischen Transformation der Rasterdaten kann für jedes Pixel im Originalbild eine neue Position im Zielbild berechnet werden. Da im Ergebnis allerdings einige Pixel mehrfach oder gar nicht belegt werden, müssen anschließend durch Auswahl- und Interpolationsverfahren Korrekturen durchgeführt werden. Bei dem *indirekten* Verfahren wird durch Umkehrung der Transformationsvorschrift für jedes Pixel im Zielbild genau ein Intensitätswert aus dem Originalbild entnommen (vgl. Abb. 3.1b).

Da bei diesen Berechnungen die Bildpositionen in der Regel nicht ganzzahlig sind, sollten die Intensitätswerte mit Hilfe der benachbarten Pixel geeignet interpoliert werden. Eine verbreitete Methode stellt die *bilineare Interpolation* dar, bei der mit vergleichsweise geringem Berechnungsaufwand ein gewichtetes Mittel bestimmt wird. Unter Berücksichtigung der ursprünglichen kontinuierlichen Intensitätswertfunktion verringert dieses Verfahren jedoch ein wenig die Bildkontraste.



**Abb. 3.1:** Anpassung der Intensitätswerte für die Vergleichsmuster

Qualitativ bessere Ergebnisse kann man mit einer Interpolation höherer Ordnung erzielen. Die *bikubische Interpolation* bestimmt als ein Verfahren zweiter Ordnung den gesuchten Intensitätswert aus 16 Pixeln in einer  $4 \times 4$  Nachbarschaft. Die Berechnungsvorschrift

$$f(i+dx, j+dy) = \sum_{m=-1}^2 \sum_{n=-1}^2 f(i+m, j+n) \cdot r(m-dx) \cdot r(dy-n) \quad (7.1)$$

mit der kubischen Gewichtungsfunktion

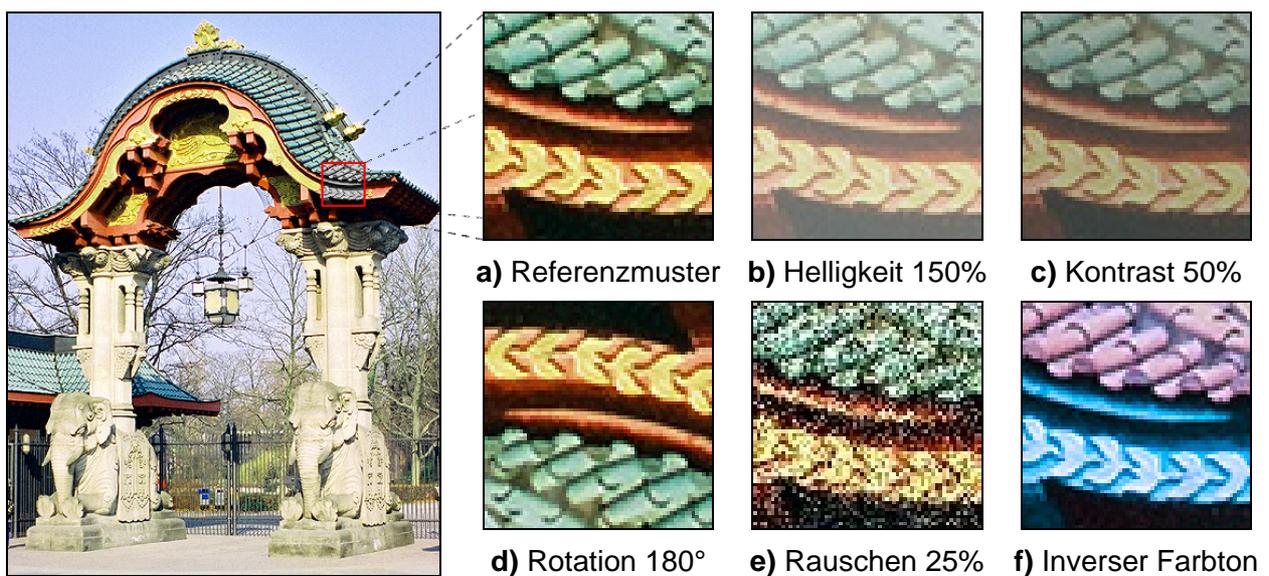
$$r(k) = \frac{1}{6} \left[ p(k+2)^3 - 4p(k+1)^3 + 6p(k)^3 - 4p(k-1)^3 \right] \quad (7.2)$$

mit  $p(k) = \begin{cases} k & \text{für } k > 0 \\ 0 & \text{sonst} \end{cases}$

definiert den interpolierten Intensitätswert und muss für jeden Farbkanal einzeln berechnet werden.

## 3.2 Ähnlichkeitsmaße

Nachfolgend sollen verschiedene Methoden untersucht werden, die eine Ähnlichkeit von zwei farbigen Bildausschnitten beschreiben. Neben der schnellen *Differenz-Korrelation* (Abschnitt 3.2.1) und der robusten *normierten Kreuz-Korrelation* (Abschnitt 3.2.2) werden auch höhere Merkmale mit Hilfe von *differentiellen Invarianten* berücksichtigt (Abschnitt 3.2.3). Ein abschließender Vergleich demonstriert die einzelnen Verwendungsmöglichkeiten im Nahbereich.



**Abb. 3.2:** Bildausschnitt vom ELEFANTENTOR des Zoologischen Gartens in Berlin zur Analyse der Korrelationsmethoden

### 3.2.1 Differenz-Korrelation

Die einfache *Summe der Differenz-Quadrate* (Sum of Squared Differences, SSD) wird häufig bei leichten Variationen der Intensitätswerte als relativ empfindliches Maß eingesetzt, weil sie einen geringen Berechnungsaufwand erfordert. Für die Bestimmung sollten zunächst die Farbwertanteile durch ihre Intensität normiert werden

$$f' = \frac{f}{r + g + b} \quad \text{für } f = r, g, b. \quad (3.1)$$

Anschließend kann für zwei Bildausschnitte  $a(x, y)$  und  $b(x, y)$  der Größe  $n \times n$  die folgende Farbdifferenz definiert werden:

$$\rho_{DK}(a, b) = \frac{1}{3n^2} \sum_{i,j=1}^n \sum_{k=R,G,B} [a'_k(i, j) - b'_k(i, j)]^2 \quad (3.2)$$

Um ein anschauliches Ähnlichkeitsmaß zu erhalten und die Angabe eines Schwellwertes zu erleichtern, sollten die Werte des Korrelations-Koeffizienten mit

$$\rho'(a, b) = \frac{1}{1 + \rho(a, b)} \quad (3.3)$$

in das Intervall  $[0..1]$  transformiert werden. Ein Beispiel für die parallele Implementierung der Differenz-Korrelation zur Beschleunigung von Bildzuordnungs-Verfahren unter Echtzeit-Bedingungen ist z.B. in KOSCHAN und RODEHORST [KOS95] beschrieben.

### 3.2.2 Normierte Kreuz-Korrelation

Im Gegensatz zu den einfachen Farbdifferenzen verhalten sich die etwas aufwändigeren statistischen Ansätze weitaus stabiler gegenüber Störungen. In der Praxis wird häufig die normierte Kreuz-Korrelation eingesetzt. Der für Farbbilder entsprechend modifizierte Korrelations-Koeffizient lautet

$$\begin{aligned} \rho_{KK}(a, b) &= \frac{\sigma_{ab}}{\sqrt{\sigma_a^2 \cdot \sigma_b^2}} \\ &= \frac{\frac{1}{3n^2} \sum_{i,j=1}^n \sum_{k=R,G,B} a_k(i, j) \cdot b_k(i, j) - \bar{a} \cdot \bar{b}}{\sqrt{\left( \frac{1}{3n^2} \sum_{i,j=1}^n \sum_{k=R,G,B} a_k(i, j)^2 - \bar{a}^2 \right) \cdot \left( \frac{1}{3n^2} \sum_{i,j=1}^n \sum_{k=R,G,B} b_k(i, j)^2 - \bar{b}^2 \right)}}, \end{aligned} \quad (3.4)$$

wobei die Schreibweise

$$\bar{a} = \frac{1}{3n^2} \sum_{i,j=1}^n \sum_{k=R,G,B} a_k(i, j)$$

den mittleren Intensitätswert für den Bildausschnitt  $a(x, y)$  bezeichnet. Die normierte Kreuz-Korrelation ist robust gegenüber *radiometrischen* Unterschieden, wie z.B. Variationen in der Helligkeit oder im Kontrast.

Im Extremfall wird durch einen homogen weißen und einen komplett schwarzen Bildausschnitt mit  $\rho_{KK} = 1$  eine optimale Übereinstimmung definiert. Genau wie die Differenz-Korrelation ist die Kreuz-Korrelation aufgrund des pixelweisen Vergleichs nicht invariant gegenüber *geometrischen* Verzerrungen, wie z.B. Rotationen in der Bildebene oder Projektionen aus verschiedenen Blickwinkeln. So ergibt ein Schachbrett-Muster, das mit einer um 90 Grad gedrehten Kopie verglichen wird, einen minimalen Korrelations-Koeffizienten von  $\rho_{KK} = 0$ . Um die im Nahbereich üblicherweise vorkommenden Deformationen besser berücksichtigen zu können, empfiehlt sich der Einsatz von höheren Merkmalen.

### 3.2.3 Differentielle Invarianten

Der Inhalt eines Bildausschnitts lässt sich auch mit Invarianten charakterisieren, die nicht von der Lage oder Drehung abhängig sind. Die Anwendung eignet sich aber hauptsächlich für Farbbilder, da bei Grauwerten die Berechnung der Merkmale bis zur dritten Ordnung sehr aufwändig und instabil ist.

Nach MINDRU u.a. [MIN01] ist ein allgemeines *Farb-Moment* für eine digitale Bildfunktion  $f = (r, g, b)^T$  in dem Bereich  $\Omega$  durch

$$M_{PQ}^{ABC} = \sum_{(x,y) \in \Omega} x^P y^Q \cdot r(x,y)^A \cdot g(x,y)^B \cdot b(x,y)^C \quad (3.5)$$

definiert und besitzt die *Ordnung*  $P + Q$  sowie den *Grad*  $A + B + C$ . Bei einer Beschränkung auf Ableitungen erster Ordnung sind diese Momente robust gegenüber Rauschen und relativ einfach zu berechnen. Die Mittelwerte der drei einzelnen Farbkanäle (0. Ordnung) und die durchschnittlichen Gradientenbeträge (1. Ordnung) können anhand der kontinuierlichen Differential-Operatoren unter Angabe der Standardabweichung  $\sigma$  leicht berechnet werden (vgl. Abschnitt 2.1.4).

Nach der Invarianten-Theorie von HILBERT [HIL90] müssen bei drei Farbkanälen aber insgesamt  $3 \times 3 - 1 = 8$  Komponenten definiert werden. Daher schlägt MONTESINOS [MON98] noch zwei zusätzliche Farbkombinationen vor, so dass sich für ein Gebiet  $\Omega$  schließlich der Merkmals-Vektor

$$\mathbf{x} = \left( \sum r, \sum g, \sum b, \sum |\nabla r|^2, \sum |\nabla g|^2, \sum |\nabla b|^2, \sum \nabla r \nabla g, \sum \nabla r \nabla b \right)^T \quad (3.6)$$

mit acht unabhängigen Komponenten ergibt. Diese Charakterisierung eines Bildausschnitts ist allerdings nicht invariant gegenüber radiometrischen Unterschieden.

Nach FINLAYSON [FIN94] kann aber das folgende *lineare Modell* verwendet werden

$$f'(x, y) = d_1 \cdot f(x, y) + d_2, \quad (3.7)$$

wobei  $d_1$  die Unterschiede im *Kontrast* und  $d_2$  in der *Helligkeit* des Bildes darstellen.

Um die Intensitäten unabhängig von den Modellparametern auszudrücken, eignet sich eine lineare Transformation der Bildinhalte auf das Intervall  $[0 \dots 1]$  mit

$$f'(x, y) = \frac{f(x, y) - f_{\min}}{f_{\max} - f_{\min}}, \quad (3.8)$$

wobei  $f_{\min}$  und  $f_{\max}$  den minimalen bzw. maximalen im Bild vorkommenden Intensitätswert bezeichnen. Der Einfluss dieser radiometrischen Normierung auf die Komponenten des Merkmals-Vektors ist in Abb. 3.3 dargestellt.

Nachdem die Merkmals-Vektoren berechnet wurden, müssen sie miteinander verglichen werden. Da jede Vektorkomponente einen eigenen Wertebereich besitzt, kann der euklidische Abstand nicht direkt für einen Vergleich herangezogen werden. Eine Möglichkeit wäre, jede Komponente auf das Intervall  $[0 \dots 1]$  zu normieren, indem die Extrema einer Komponente in allen verfügbaren Merkmals-Vektoren gesucht werden.

Da aber neben dem Wertebereich auch die statistische Verteilung variieren kann, empfiehlt sich für den Vergleich der Vektoren  $\mathbf{x}$  und  $\mathbf{y}$  der *Mahalanobis-Abstand*

$$\rho_{INV}(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})^T \mathbf{K}^{-1} (\mathbf{x} - \mathbf{y})}, \quad (3.9)$$

wobei  $\mathbf{K}$  die *Kovarianz-Matrix* der  $n$  Vektorkomponenten bezeichnet.

Dieses Abstandsmaß folgt der  $\chi^2$ -Verteilung und entspricht dem euklidischen Abstand, wenn  $\mathbf{K}$  die Einheits-Matrix  $\mathbf{I}$  ist. Um einen anschaulichen Korrelations-Koeffizienten zu erhalten, sollte auch hier der Wertebereich mit Gleichung (3.3) transformiert werden.

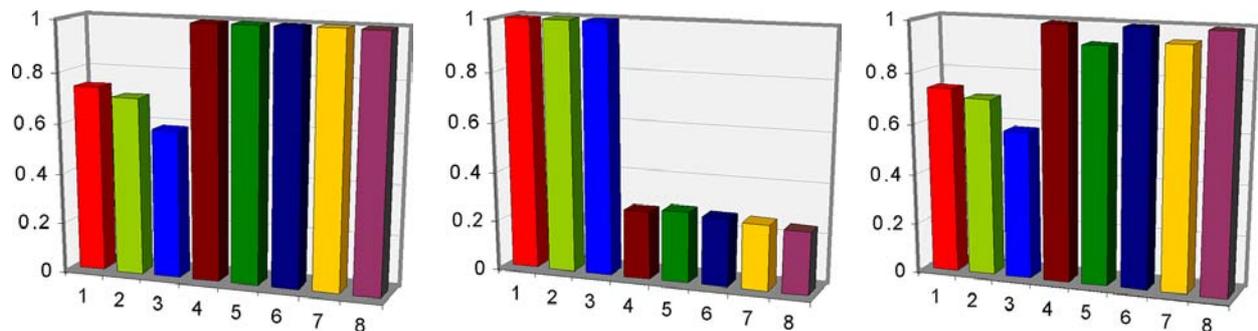
Eine wesentliche Aufgabe bei der Bestimmung des Mahalanobis-Abstands besteht in der Schätzung der zur Hauptachse symmetrischen Kovarianz-Matrix.

$$\mathbf{K} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}, \quad (3.10)$$

wobei die *Kovarianz* des Vektorkomponenten-Paares  $(i, j)$  durch

$$\sigma_{ij} = \frac{1}{m} \sum_{k=1}^m \mathbf{x}_{ik} \mathbf{y}_{jk} - \bar{\mathbf{x}}_i \cdot \bar{\mathbf{y}}_j \quad \text{für } i, j = 1, \dots, n \quad (3.11)$$

mit  $n = 8$  definiert ist. Dabei bezeichnen  $\bar{\mathbf{x}}$  und  $\bar{\mathbf{y}}$  die Mittelwerte in den Komponenten  $i$  bzw.  $j$  bei insgesamt  $m$  verfügbaren Merkmals-Vektoren.



a) Merkmals-Vektor für das Referenzmuster in Abb. 3.2a    b) Vektor für das aufgehellte Muster in Abb. 3.2b    c) Vektor nach der Normierung von Abb. 3.2b

**Abb. 3.3:** Radiometrische Invarianz der Merkmals-Vektoren durch Normierung. Die einzelnen Komponenten sind für die Darstellung auf den Betrag 1 skaliert worden.

### 3.2.4 Vergleich der Methoden

Um die Leistungsfähigkeit der vorgestellten Ähnlichkeitsmaße bewerten zu können, sind in Tab. 3.1 die Korrelations-Ergebnisse mit den Testbildern aus Abb. 3.2 zusammengefasst. Bis auf den Vergleich mit dem inversen Farbton-Muster ( $a \leftrightarrow f$ ) sollten alle Werte dicht bei 1 liegen.

Es zeigt sich, dass die normierte Kreuz-Korrelation bei radiometrischen Unterschieden und insbesondere bei Variationen des Kontrastes ( $a \leftrightarrow c$ ) weitaus robuster als die Differenz-Korrelation ist. Auch für die Unterscheidung von Farbtönen werden wesentlich deutlichere Ergebnisse produziert. Insgesamt liefert die normierte Kreuz-Korrelation hervorragende Ergebnisse, sofern die geometrischen Unterschiede zwischen den Bildausschnitten verhältnismäßig gering sind.

Die experimentellen Versuche mit den differentiellen Invarianten haben gezeigt, dass aufgrund der geometrischen Invarianz die Ähnlichkeit von zwei Bildausschnitten nicht präzise beschrieben werden kann. Sie eignen sich allerdings dazu, durch einen Ausschluss von unverträglichen Merkmalen die Anzahl möglicher Korrespondenz-Kandidaten zu reduzieren und dadurch das Verfahren zu beschleunigen (vgl. Tab. 3.3). Die differentiellen Invarianten sollten daher als notwendiges, aber nicht unbedingt hinreichendes Kriterium angesehen werden.

Ähnlichkeit	Testbilder				
	$a \leftrightarrow b$	$a \leftrightarrow c$	$a \leftrightarrow d$	$a \leftrightarrow e$	$a \leftrightarrow f$
$\rho_{DK}$	0,92	0,70	0,89	0,87	0,75
$\rho_{KK}$	0,94	0,94	0,65	0,92	0,44
$\rho_{INV}$	0,98	0,98	1,00	0,76	0,65

**Tab. 3.1:** Vergleich der Korrelations-Koeffizienten

### 3.3 Geometrische Anpassung

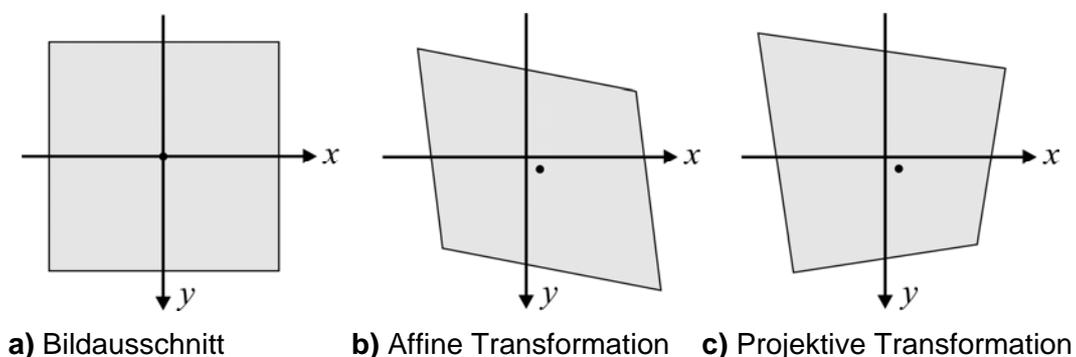
Da die vorgestellten Ähnlichkeitsmaße für die im Nahbereich vorkommenden Deformationen ungeeignet sind, sollten vorher die Unterschiede bei den Verzerrungen kompensiert werden, d.h. es wird eine geometrische Transformation gesucht, die beide Bildmuster möglichst gut in Übereinstimmung bringt.

In der Photogrammetrie wird dieses Problem mit der Kleinste-Quadrate-Bildzuordnung (*Least Squares Matching*) gelöst [HEI90, FÖR85, ACK84], wobei sich der folgende Abschnitt im Wesentlichen auf das geometrische Modell konzentriert (Abschnitt 3.3.1). Nach einigen Anmerkungen zur Implementierung des Verfahrens (Abschnitt 3.3.2) wird schließlich die Leistungsfähigkeit der geometrischen Anpassung anhand einiger Beispiele demonstriert (Abschnitt 3.3.3).

#### 3.3.1 Deformations-Modell

Die Verzerrungen können bei Objektoberflächen, die stückweise als planar angenommen werden, mit Hilfe einer Bildtransformation modelliert werden. Die Deformation eines abgebildeten Flächenstücks kann dabei am besten durch die acht Parameter einer projektiven Transformation beschrieben werden (siehe Abb. 3.4). In der Praxis wird das *projektive Modell* aber nur selten eingesetzt, da es hochgradig nicht-linear und daher schwer zu berechnen ist.

SZELISKI [SZE94] versucht z.B., die projektive Deformation mit Hilfe einer LEVENBERG-MARQUARDT-Optimierung zu rekonstruieren. Die Ergebnisse bei der Erstellung von Bildmosaiken sind auch recht eindrucksvoll, jedoch werden hier sehr große Bildbereiche vorausgesetzt. Für eine Anwendung im Nahbereich sollten die Flächenstücke im Verhältnis zum gesamten Bild möglichst klein sein. In diesem Fall lässt sich eine projektive Transformation durch die sechs Parameter einer affinen Transformation approximieren [MAN92].



**Abb. 3.4:** Geometrische Deformations-Modelle

Bei dem *affinen Modell* lässt sich die Transformation eines Pixels an der Stelle  $\mathbf{x} = (x, y)^T$  folgendermaßen darstellen:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d_5 \\ d_6 \end{bmatrix} \quad \text{bzw.} \quad \mathbf{x}' = \mathbf{D}\mathbf{x} + \mathbf{d}, \quad (3.12)$$

wobei  $\mathbf{D}$  die *Deformations-Matrix* und  $\mathbf{d}$  den *Verschiebungs-Vektor* bezeichnen. Für die Berechnung der Transformations-Parameter eignet sich allerdings ein leicht verändertes Modell, wobei ein Punkt  $\mathbf{x}$  des ersten Bildes  $f$  sich an die Stelle  $\mathbf{x}' = [\mathbf{I} + \mathbf{D}]\mathbf{x} + \mathbf{d}$  im zweiten Bild  $g$  bewegt. Diese Modifikation hat den Vorteil, dass für eine identische Abbildung alle Transformations-Parameter 0 sind.

Die Aufgabe der geometrischen Anpassung besteht nun darin, die sechs Deformationsparameter von  $\mathbf{D}$  und  $\mathbf{d}$  zu finden, so dass die Differenzen in den Intensitätswerten der beiden Bildausschnitte  $f(\mathbf{x})$  und  $g(\mathbf{x}')$  minimiert werden

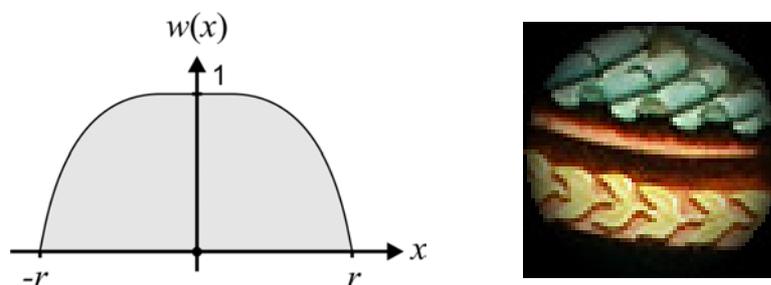
$$\min_{\mathbf{D}, \mathbf{d}} \sum_{\mathbf{x} \in \Omega} \left[ f(\mathbf{x}) - g([\mathbf{I} + \mathbf{D}]\mathbf{x} + \mathbf{d}) \right]^2 w(\mathbf{x}), \quad (3.13)$$

wobei die Koordinaten für das Gebiet  $\Omega$  mit dem Radius  $r$  bezüglich der Bildmitte gemessen werden und  $w$  eine optionale Gewichtungsfunktion darstellt. Bei der Verwendung einer GAUSS-ähnlichen Funktion oder der symmetrischen Funktion nach MANMATHA und OLIENSIS [MAN92]

$$w(\mathbf{x}) = \begin{cases} \cos\left(\frac{\mathbf{x}^T \mathbf{x}}{\sigma^2}\right) & \text{für } -\frac{\pi}{2} \leq \left| \frac{\mathbf{x}}{\sigma} \right| \leq \frac{\pi}{2} \\ 0 & \text{sonst} \end{cases} \quad \text{mit } \sigma = \frac{r}{\sqrt{2\pi}}, \quad (3.14)$$

kann der zentrale Bereich um den Interest-Punkt betont werden (siehe Abb. 3.5).

Falls die Position eines Interest-Punktes mit Subpixel-Genauigkeit bestimmt werden konnte, dann kann auf die Berechnung des Verschiebungs-Vektors  $\mathbf{d}$  verzichtet werden (*Reduziertes Deformations-Modell*).



a) Funktionsprofil in  $x$ -Richtung b) Gewichtetes Referenzmuster

**Abb. 3.5:** Einfluss der Gewichtungsfunktion

### 3.3.2 Kleinste-Quadrate-Bildzuordnung

Obwohl das Problem in Gleichung (3.14) eigentlich nicht linear ist, können die gesuchten Parameter für den Bildbereich  $\Omega$  durch ein linearisiertes Gleichungssystem der Form

$$\mathbf{A} \mathbf{z} = \mathbf{b} \quad (3.15)$$

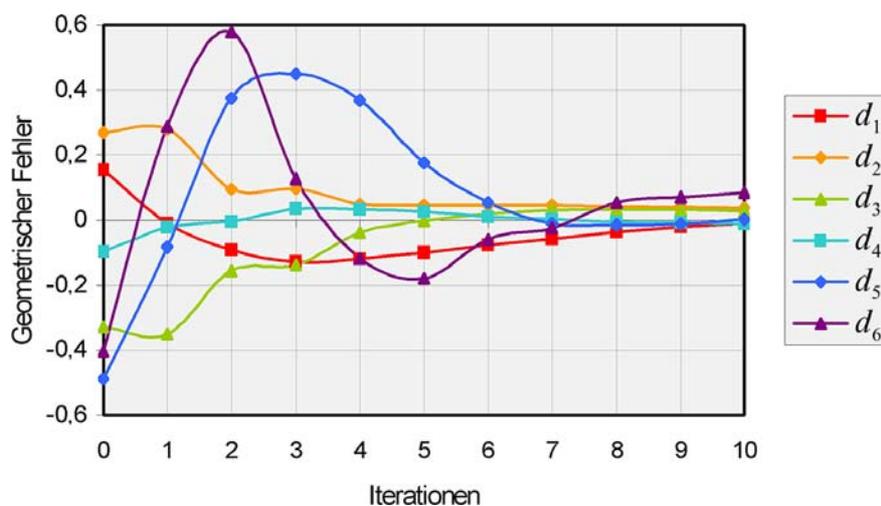
approximiert werden, wobei der Vektor

$$\mathbf{b} = (f - g) \cdot \left( \sum \hat{f}_x x, \sum \hat{f}_x y, \sum \hat{f}_y x, \sum \hat{f}_y y, \sum \hat{f}_x, \sum \hat{f}_y \right)^T \quad (3.16)$$

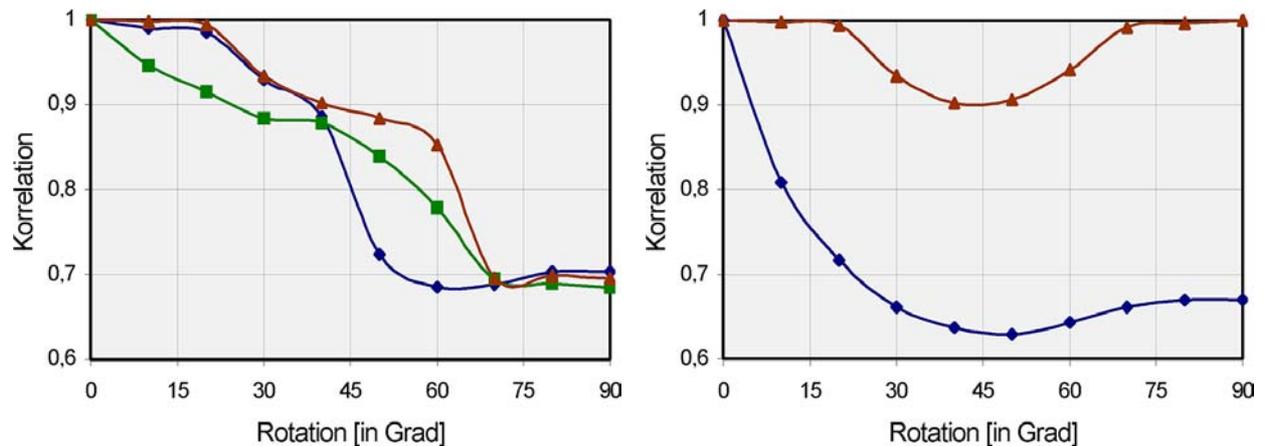
von den Differenzen und der Addition  $\hat{f} = f + g$  der beiden Ausschnitte abhängt und die symmetrische Design-Matrix  $\mathbf{A}$  durch

$$\mathbf{A} = \begin{bmatrix} \sum \hat{f}_x^2 x^2 & \sum \hat{f}_x^2 xy & \sum \hat{f}_x \hat{f}_y x^2 & \sum \hat{f}_x \hat{f}_y xy & \sum \hat{f}_x^2 x & \sum \hat{f}_x \hat{f}_y x \\ \sum \hat{f}_x^2 xy & \sum \hat{f}_x^2 y^2 & \sum \hat{f}_x \hat{f}_y xy & \sum \hat{f}_x \hat{f}_y y^2 & \sum \hat{f}_x^2 y & \sum \hat{f}_x \hat{f}_y y \\ \sum \hat{f}_x \hat{f}_y x^2 & \sum \hat{f}_x \hat{f}_y xy & \sum \hat{f}_y^2 x^2 & \sum \hat{f}_y^2 xy & \sum \hat{f}_x \hat{f}_y x & \sum \hat{f}_y^2 x \\ \sum \hat{f}_x \hat{f}_y xy & \sum \hat{f}_x \hat{f}_y y^2 & \sum \hat{f}_y^2 xy & \sum \hat{f}_y^2 y^2 & \sum \hat{f}_x \hat{f}_y y & \sum \hat{f}_y^2 y \\ \sum \hat{f}_x^2 x & \sum \hat{f}_x^2 y & \sum \hat{f}_x \hat{f}_y x & \sum \hat{f}_x \hat{f}_y y & \sum \hat{f}_x^2 & \sum \hat{f}_x \hat{f}_y \\ \sum \hat{f}_x \hat{f}_y x & \sum \hat{f}_x \hat{f}_y y & \sum \hat{f}_y^2 x & \sum \hat{f}_y^2 y & \sum \hat{f}_x \hat{f}_y & \sum \hat{f}_y^2 \end{bmatrix} \quad (3.17)$$

definiert ist. Der gesuchte Lösungsvektor  $\mathbf{z} = (d_1, \dots, d_4, d_5, d_6)^T$  kann z.B. nach der Methode der kleinsten Quadrate numerisch bestimmt werden (vgl. Anhang A.2.2). Da die Gleichung (3.15) durch Linearisierung mit TAYLOR entwickelt wurde, müssen die Parameter allerdings noch iterativ verbessert werden. Das Konvergenzverhalten bei der geometrischen Anpassung ist in Abb. 3.6 am Beispiel eines um 30 Grad rotierten Bildes dargestellt.



**Abb. 3.6:** Konvergenz des geometrischen Fehlers bei der iterativen Optimierung der affinen Deformationsparameter  $d_1, \dots, d_6$  über zehn Iterationen



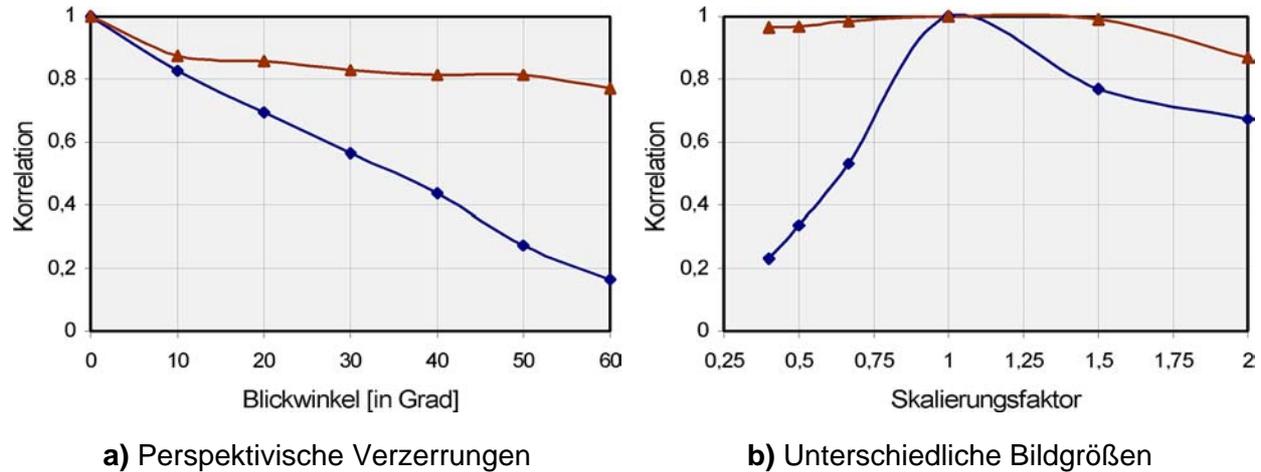
**a)** Adaptive Glättung mit  $\sigma = 2,5 \dots 0,5$  ( $\blacktriangle$ ) im Vergleich zu konstanten Standardabweichungen  $\sigma = 1,0$  ( $\blacklozenge$ ) und  $\sigma = 3,0$  ( $\blacksquare$ ) **b)** Geometrische Anpassung mit rotierten Referenzbildern ( $\blacktriangle$ ) im Vergleich zur Kreuzkorrelation ohne Deformations-Modell ( $\blacklozenge$ )

**Abb. 3.7:** Ergebnis der geometrische Anpassung bei rotierten Bildausschnitten

Um die Konvergenz des Verfahrens zu optimieren, wurde eine *adaptive Glättung* eingeführt. Der Wert der Standardabweichung  $\sigma$  für die partiellen Ableitungen  $\hat{f}_x$  und  $\hat{f}_y$  beeinflusst die Stabilität der Berechnung. Mit zunehmender Glättung werden zwar größere geometrische Unterschiede besser, dafür aber auch kleine Differenzen schlechter rekonstruiert. Es empfiehlt sich daher, mit einer größeren Glättung zu beginnen und diese im Laufe der Berechnung sukzessive zu reduzieren. Die Verbesserung durch die adaptive Glättung wird in Abb. 3.7 am Beispiel einer Bildrotation illustriert.

Die Resultate zeigen aber auch, dass nur eine Drehung von maximal 60 Grad je Richtung bestimmt werden kann. Um diesen Bereich für Bilder zu erweitern, die sowohl im Hoch- als auch Querformat vorliegen, kann das Verfahren viermal mit einem jeweils um 90 Grad gedrehten Referenzbild ausgeführt werden. Neben der Rotation sind aber noch weitere Deformationen relevant. Die Verbesserungen der Korrelationsergebnisse durch eine geometrische Anpassung von perspektivischen Verzerrungen und unterschiedlichen Bildgrößen sind daher in Abb. 3.8 dargestellt.

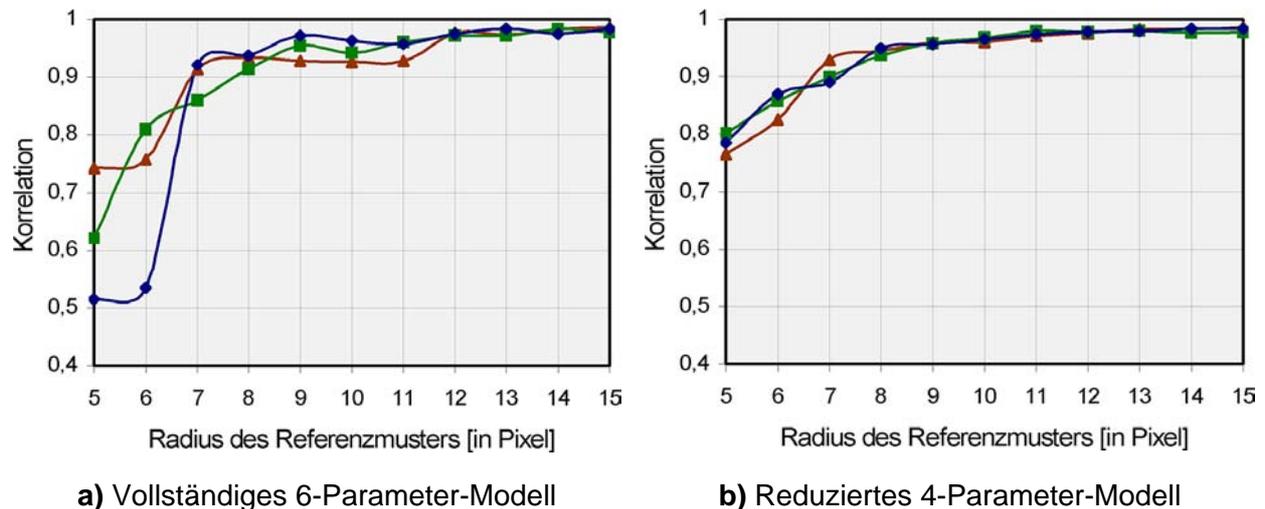
Da insbesondere bei Unterschieden in der Skalierung die Richtung des Vergleichs eine Rolle spielt, werden jeweils beide Bildausschnitte einmal als Referenzmuster definiert und anschließend das bessere Korrelationsergebnis ausgewählt. Durch die zehn Iterationen mit vier gedrehten Referenzmustern und dem symmetrischen Vergleich erfordert die geometrische Anpassung allerdings einen hohen Berechnungsaufwand (vgl. Tab. 3.3).



**Abb. 3.8:** Ähnlichkeit von deformierten Bildausschnitten vor (◆) und nach (▲) der Anpassung

Die Qualität der geometrischen Anpassung hängt insbesondere von der Größe der verwendeten Vergleichsmuster ab. Die Untersuchungen haben gezeigt, dass für die Bildausschnitte ein Radius von mindestens sieben Pixeln gewählt werden sollte (vgl. Abb. 3.9). Wenn durch die Subpixel-Information der Interest-Punkte auf die Berechnung des Verschiebungsvektors  $\mathbf{d}$  verzichtet werden kann, dann zeigt sich besonders bei den kleinen Bildausschnitten eine Stabilisierung der Ergebnisse.

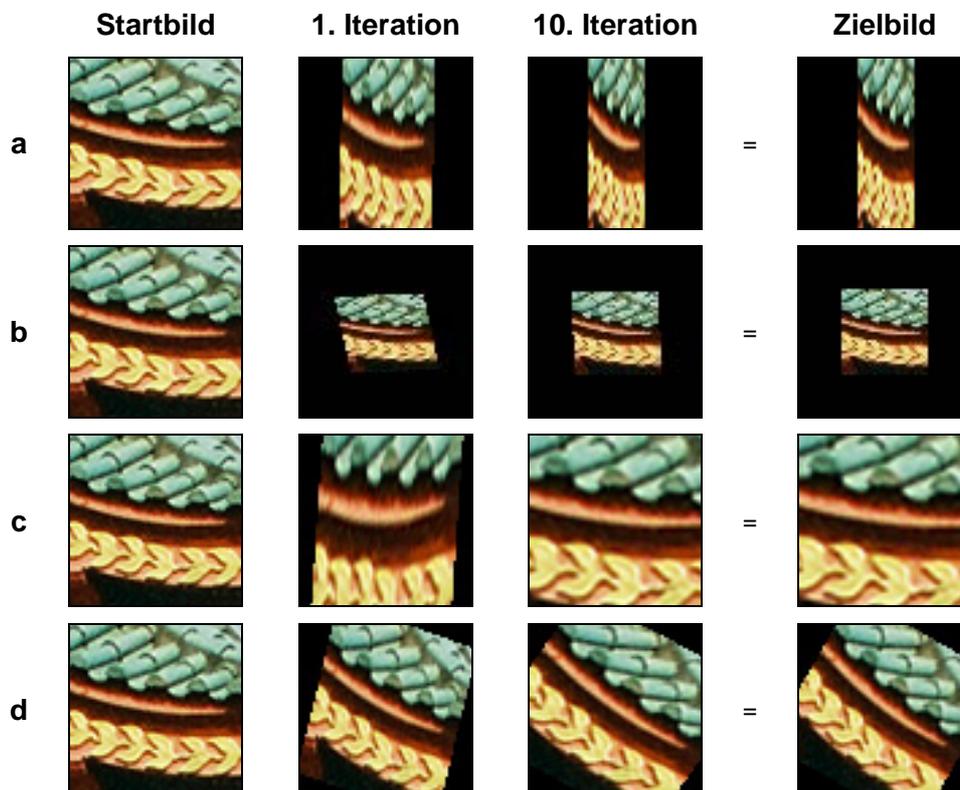
Um die Kleinste-Quadrate-Bildzuordnung für Farbbilder zu erweitern, kann die Gleichung (3.15) analog zum Abschnitt 2.3.3 relativ einfach modifiziert werden.



**Abb. 3.9:** Stabilität der geometrischen Anpassung in Abhängigkeit von der Größe der Vergleichsmuster am Beispiel einer Rotation (▲), einer Perspektiv-Verzerrung (■) und einer Skalierung (◆)

### 3.3.3 Experimentelle Ergebnisse

Um die Wirkungsweise und Leistungsfähigkeit der geometrischen Anpassung zu veranschaulichen, sind verschiedene Beispiele für die Berechnung der Deformations-Matrizen in Abb. 3.10 visualisiert und die zugehörigen numerischen Ergebnisse in Tab. 3.2 zusammengefasst.



**Abb. 3.10:** Beispiele für die geometrische Anpassung eines Bildausschnitts an ein deformiertes Vergleichsmuster in zehn Iterationen

Bsp.	Wahre Transformation	Rekonstruierte Transformation	$\rho_{KK}$	Ergebnis mit Gewichtung $w$	$\rho_{KK}$	Ergebnis bei 25% Rauschen	$\rho_{KK}$
a	$\begin{bmatrix} 3,00 & 0,00 \\ 0,00 & 0,67 \end{bmatrix}$	$\begin{bmatrix} 2,99 & -0,02 \\ -0,01 & 0,66 \end{bmatrix}$	1,00	$\begin{bmatrix} 2,94 & 0,02 \\ 0,01 & 0,69 \end{bmatrix}$	0,98	$\begin{bmatrix} 3,00 & 0,02 \\ 0,00 & 0,67 \end{bmatrix}$	0,93
b	$\begin{bmatrix} 0,67 & 0,00 \\ 0,00 & 0,67 \end{bmatrix}$	$\begin{bmatrix} 0,68 & -0,01 \\ 0,00 & 0,67 \end{bmatrix}$	1,00	$\begin{bmatrix} 0,66 & 0,01 \\ 0,00 & 0,68 \end{bmatrix}$	1,00	$\begin{bmatrix} 0,79 & 0,08 \\ 0,01 & 0,68 \end{bmatrix}$	0,92
c	$\begin{bmatrix} 2,00 & 0,00 \\ 0,00 & 2,00 \end{bmatrix}$	$\begin{bmatrix} 1,99 & -0,01 \\ 0,00 & 2,00 \end{bmatrix}$	0,99	$\begin{bmatrix} 2,00 & 0,02 \\ 0,00 & 1,95 \end{bmatrix}$	0,98	$\begin{bmatrix} 2,04 & 0,03 \\ 0,00 & 1,94 \end{bmatrix}$	0,90
d	$\begin{bmatrix} 0,87 & 0,50 \\ -0,50 & 0,87 \end{bmatrix}$	$\begin{bmatrix} 0,73 & 0,52 \\ -0,46 & 0,84 \end{bmatrix}$	0,94	$\begin{bmatrix} 0,79 & 0,46 \\ -0,49 & 0,89 \end{bmatrix}$	0,98	$\begin{bmatrix} 0,82 & 0,46 \\ -0,52 & 0,89 \end{bmatrix}$	0,93

**Tab. 3.2:** Numerischer Vergleich der berechneten Deformations-Matrizen  $\mathbf{D}$  für die visualisierten Beispiele in Abb. 3.10

Dabei wird deutlich, dass der Einsatz der Gewichtungsfunktion  $w$  besonders die Ergebnisse bei gedrehten Bildausschnitten verbessert. Selbst bei verrauschten Bilddaten (vgl. Abb. 3.2e) zeigt die Kleinste-Quadrate-Bildzuordnung noch stabile Resultate. Da die Bildausschnitte sich hier aufgrund der Störungen unterscheiden, kann die Ähnlichkeit nach der geometrischen Anpassung nicht den optimalen Wert 1,0 erreichen.

### 3.4 Verfahren zur Korrespondenz-Analyse

Nachdem verschiedene Techniken und Methoden für den Ähnlichkeitsvergleich von zwei Bildausschnitten diskutiert wurden, soll nun ein optimiertes Verfahren zur merkmalsbasierten Korrespondenz-Analyse für den Nahbereich vorgestellt werden (vgl. Abb. 3.11). Das Prinzip beruht auf einer zweckmäßigen Kombination der jeweiligen positiven Eigenschaften.

Zunächst werden mit Hilfe der *differentiellen Invarianten* (Abschnitt 3.2.3) nur Kandidaten mit verträglichen Merkmalen ausgewählt ( $\rho_{INV} > 0,75$ ). Diese Vorauswahl beschleunigt die nachfolgende Korrespondenz-Analyse und reduziert falsche Zuordnungen (vgl. Tab. 3.3). Anschließend wird mit einer *geometrischen Anpassung* (Abschnitt 3.3) versucht, die zu vergleichenden Bildmuster möglichst gut in Übereinstimmung zu bringen und somit die Unterschiede bei den Verzerrungen zu kompensieren.

Das Ähnlichkeitsmaß für die ausgewählten Bildfenster liefert schließlich die radiometrisch invariante *normierte Kreuz-Korrelation* (Abschnitt 3.2.2). Danach werden die beiden Bildausschnitte mit dem höchsten Korrelations-Koeffizienten  $\rho_{KK}$  einander zugeordnet.



Abb. 3.11: Optimiertes Verfahren zur merkmalsbasierten Korrespondenz-Analyse

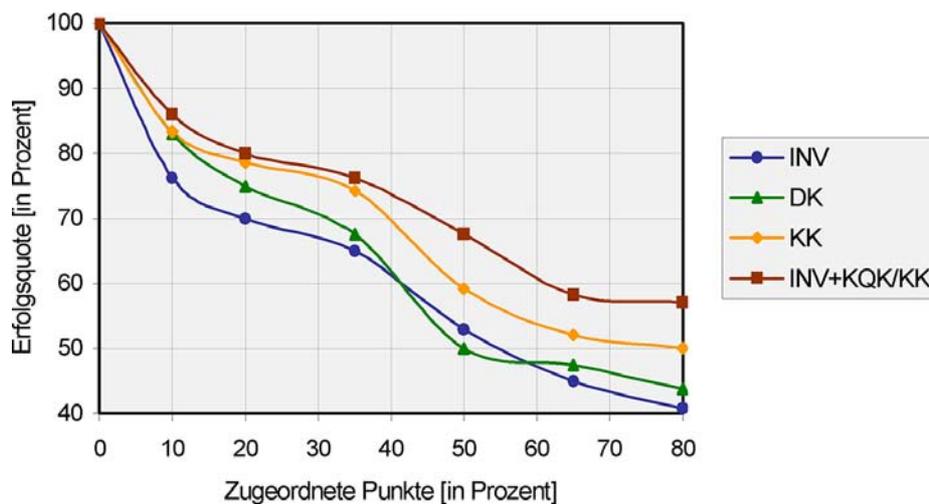
Korrelations-Verfahren	CPU [in Sek.]
Differenz-Korrelation (DK)	2,19
Normierte Kreuz-Korrelation (KK)	3,26
Differentielle Invarianten (INV)	6,90
Kleinste-Quadrate-Bildzuordnung (KQK/KK)	532,56
Kombiniertes Verfahren (INV+KQK/KK)	273,18

**Tab. 3.3:** Berechnungsaufwand für die Korrespondenz-Analyse bei 100 Punktmerkmalen

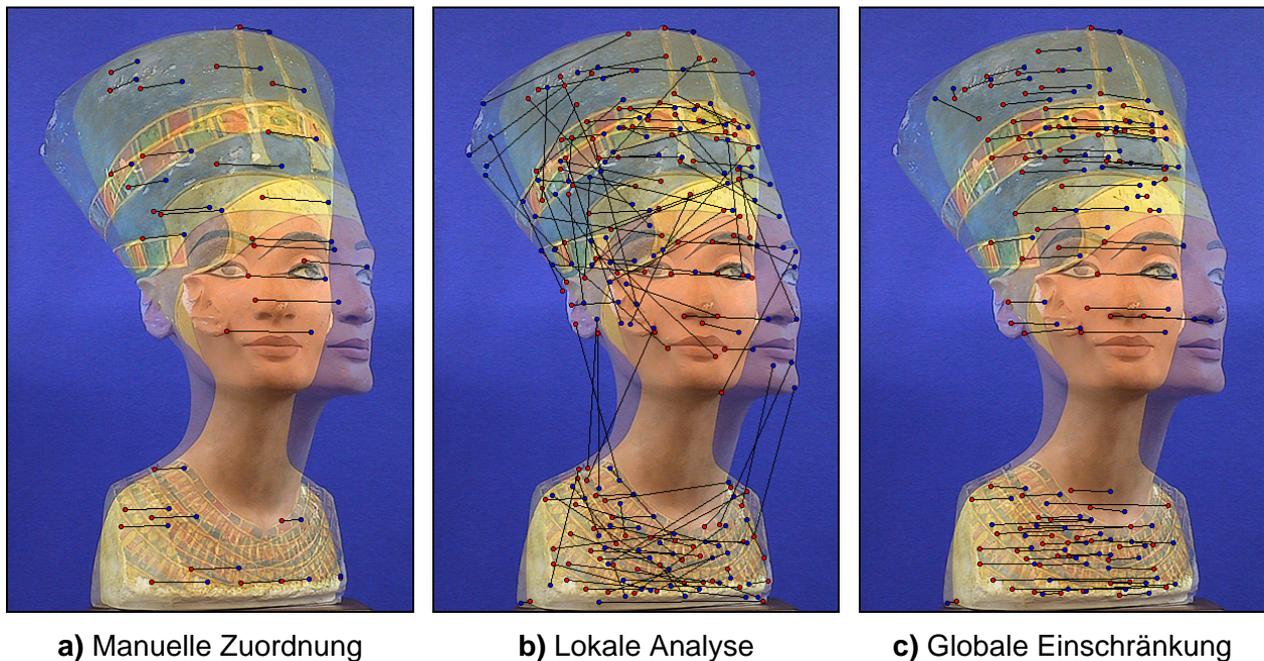
Die Leistungsfähigkeit dieser kombinierten Korrelations-Methode im Vergleich zu den Standardverfahren ist in Abb. 3.12 am Beispiel der NOFRETETE dargestellt.

Trotz dieser intensiven Bemühungen ist das optimierte Verfahren noch nicht in der Lage, zufrieden stellende Punkt-Korrespondenzen zu bestimmen. Zur Demonstration sind in Abb. 3.13b zwei Bilder der NOFRETETE-Sequenz überlagert und die homologen Punkte mit einer Linie verbunden worden. Das dargestellte Problem wird hauptsächlich durch die Beschränkung auf eine lokale Analyse mit relativ kleinen Bildausschnitten verursacht.

Um zu veranschaulichen, dass die fehlerhaften Zuordnungen nicht durch das Ähnlichkeitsmaß verursacht werden, sind in Abb. 3.14 einige korrespondierende Bildausschnitte mit ihren charakteristischen Eigenschaften visualisiert.

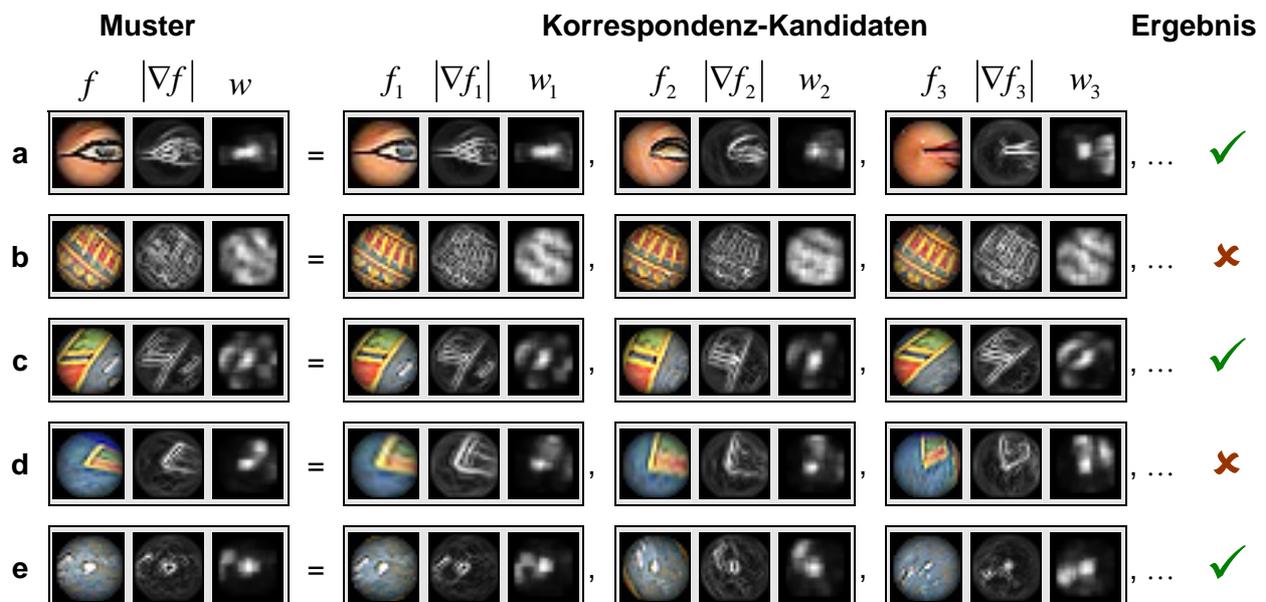


**Abb. 3.12:** Ergebnis der Korrespondenz-Analyse mit den differentiellen Invarianten (INV), der Differenz- (DK) und der Kreuz-Korrelation (KK) sowie dem kombinierten Verfahren mit Hilfe der geometrischen Anpassung (KQK)



**Abb. 3.13:** Überlagerte Stereo-Bildpaare mit interaktiv gemessenen homologen Bildpunkten [KUZ02] im Vergleich zu automatisch berechneten Korrespondenzen

Die Darstellung verdeutlicht, dass bei lokaler Betrachtung weder den Intensitätswerten  $f$  noch den Gradientenbeträgen  $|\nabla f|$  oder Punktstärken  $w$  eindeutig zu entnehmen ist, ob die entsprechende Zuordnung korrekt oder falsch ist. Aus diesem Grund ist es wichtig, das Verfahren durch zusätzliche globale Bedingungen zu unterstützen (vgl. Abb. 3.13c).



**Abb. 3.14:** Problem einer lokalen Korrespondenz-Analyse. Die Kandidaten sind mit abnehmender Ähnlichkeit sortiert, so dass die beste Übereinstimmung links dargestellt ist.

## 3.5 Bildzuordnungs-Strategien

Für die Steuerung der Korrespondenz-Analyse existieren je nach Vorwissen die unterschiedlichsten Lösungsansätze [KOS91]. Nach der Auswahl geeigneter Annahmen und Einschränkungen (Abschnitt 3.5.1) wird eine hierarchische Zuordnungs-Strategie mit Hilfe einer Auflösungs-Pyramide vorgestellt (Abschnitt 3.5.2).

### 3.5.1 Annahmen und Einschränkungen

Die Komplexität des Korrespondenzproblems wächst mit der Anzahl der vorhandenen Bildelemente. Durch Vorwissen oder Einschränkungen können Mehrdeutigkeiten bei der Zuordnung reduziert werden. Im Folgenden sollen einige Beispiele diskutiert werden.

- **Suchraum:** Die Begrenzung des Suchraums auf einen kleinen Bildausschnitt ist grundsätzlich effektiv. Üblicherweise werden die folgenden Annahmen verwendet:
  - **Epipolar-Bedingung:** Ein Bildelement kann nur mit einem Kandidaten korrespondieren, wenn dieser auf der zugehörigen Epipolar-Geraden liegt (vgl. Abschnitt 4.3.1). Die Suche nach einem Bildelement reduziert sich damit von einem zweidimensionalen auf ein eindimensionales Problem. Da die Aufnahmeconfiguration a priori nicht bekannt ist und die Epipolar-Geometrie erst durch die Korrespondenz-Analyse rekonstruiert werden soll, müssen die Bildelemente zunächst im gesamten Bild gesucht werden.
  - **Disparitätslimit:** Als weitere Einschränkung nimmt man zumeist an, dass der euklidische Abstand zwischen den homologen Punkten (*Disparität*) sehr klein ist und die Länge des Suchweges eine bestimmte Schwelle nicht überschreitet. Im Mittelpunkt dieser Arbeit steht der Nahbereich mit freier Aufnahmeanordnung und möglicherweise un stetigen Objekten, wo diese Annahme nicht zutreffend ist.
  - **Merkmalsreduktion:** Eine Begrenzung des Suchraums ist aber auch ohne Vorwissen über die Aufnahmeanordnung und die Objekte der Szene durch eine *Grob-zu-fein-Strategie* möglich (vgl. Abschnitt 3.5.2). Mit Hilfe einer Auflösungs-pyramide wird die Anzahl der möglichen Kandidaten reduziert und eine grobe Schätzung berechnet, die in den höheren Auflösungs-niveaus weiter verbessert wird.
- **Ordnungsbedingungen:** Durch die Festlegung einer Korrespondenz können Wechselwirkungen mit anderen Korrespondenzen entstehen:
  - **Eindeutigkeit:** In der Regel kann ein Bildelement nur mit genau einem Kandidaten korrespondieren. Eine Ausnahme bilden mögliche monokulare Verdeckungen, die allerdings recht selten sind.

- **Kontinuität:** Wenn sich die Tiefe entlang von Oberflächen nur stetig ändert, variieren die Disparitäten benachbarter Bildelemente ebenfalls nur geringfügig. Da dies bei unstetigen Objekten im Nahbereich und an Objektgrenzen aufgrund von Tiefensprüngen nicht sichergestellt werden kann, und die Seltenheit der extrahierten Bildelemente zu großen Abständen im Bild führt, kann die Annahme der Kontinuität nicht eingesetzt werden.
- **Reihenfolge:** Bildelemente, die auf einer Epipolar-Geraden liegen, werden in genau derselben Reihenfolge auf der entsprechenden Epipolar-Geraden abgebildet, wenn sich alle Objekte einer Szene ungefähr in der gleichen Entfernung befinden. Aufgrund der möglichen Tiefendifferenz im Nahbereich kann diese Annahme ebenfalls nicht getroffen werden.

Die Verwendung solch einschränkender Annahmen kann zwar die Mehrdeutigkeiten bei der Korrespondenz-Analyse reduzieren, aber nicht garantieren, dass weder positiv falsche noch negativ falsche Zuordnungen getroffen werden. Die Lösung ist also nicht notwendigerweise korrekt, sondern lediglich plausibel.

### 3.5.2 Hierarchischer Ansatz

Um die Komplexität des Korrespondenzproblems zu reduzieren, eignet sich ein hierarchisches Vorgehen (siehe Abb. 3.15). Die Steigerung der Zuverlässigkeit und die Beschleunigung der Bildzuordnung bei der Verwendung von Bild-Pyramiden ist z.B. in KOSCHAN, RODEHORST und SPILLER [KOS96] dargestellt.

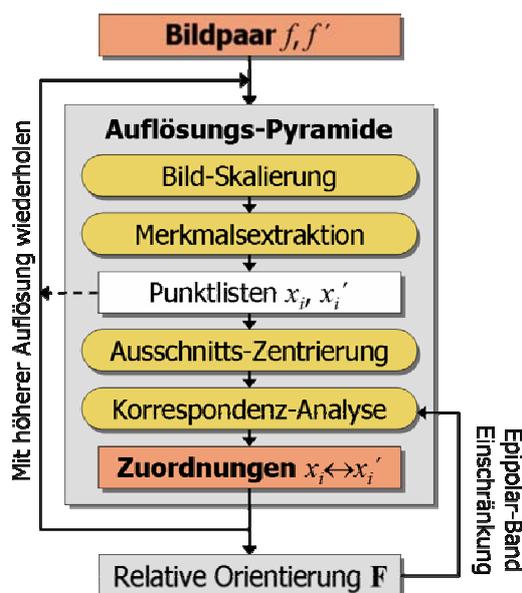
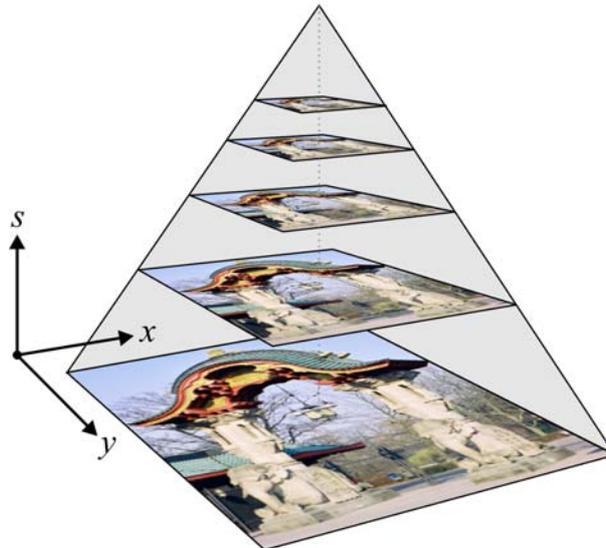


Abb. 3.15: Hierarchische Bildzuordnungs-Strategie



**Abb. 3.16:** Räumliches Modell der Auflösungs-Pyramide

Für die Bildrepräsentation wird eine *Auflösungs-Pyramide* eingesetzt, wobei das Ausgangsbild durch Faltung mit der GAUSS-Funktion geglättet (vgl. Abschnitt 2.1.3) und mit einer Unterabtastung iterativ in der Auflösung reduziert wird (siehe Abb. 3.16).

Bei der *Grob-zu-fein-Strategie* werden auf der obersten Ebene mit dem größten Auflösungs-niveau nur relativ wenige Merkmale extrahiert und die Zuordnung gestaltet sich wesentlich einfacher. Die Anzahl der Pyramidenstufen muss dabei so gewählt werden, dass noch genügend Merkmale für das Orientierungsverfahren (vgl. Abschnitt 4.3) extrahiert werden können. Mit den Ergebnissen der relativen Orientierung kann der Suchraum für die tiefere Ebene mit entsprechend höherer Auflösung eingeschränkt werden, z.B. durch eine Begrenzung mit dem *Epipolar-Band* (vgl. Abschnitt 4.3.2).

## 3.6 Experimentelle Ergebnisse

Um die Leistungsfähigkeit der vorgestellten merkmalsbasierten Korrespondenz-Analyse zu demonstrieren, wurde neben der kompakten Büste der NOFRETETE (siehe Abb. 3.13c) auch ein komplexes Objekt mit einer APS Spiegelreflex-Kamera aufgenommen. Obwohl die goldglänzende Oberfläche der VIKTORIA-Statue sehr un stetig ist und das Farbpotential des Verfahrens nicht ausnutzt, ist die Mehrzahl der Interest-Punkte korrekt zugeordnet worden (vgl. Abb. 3.17).

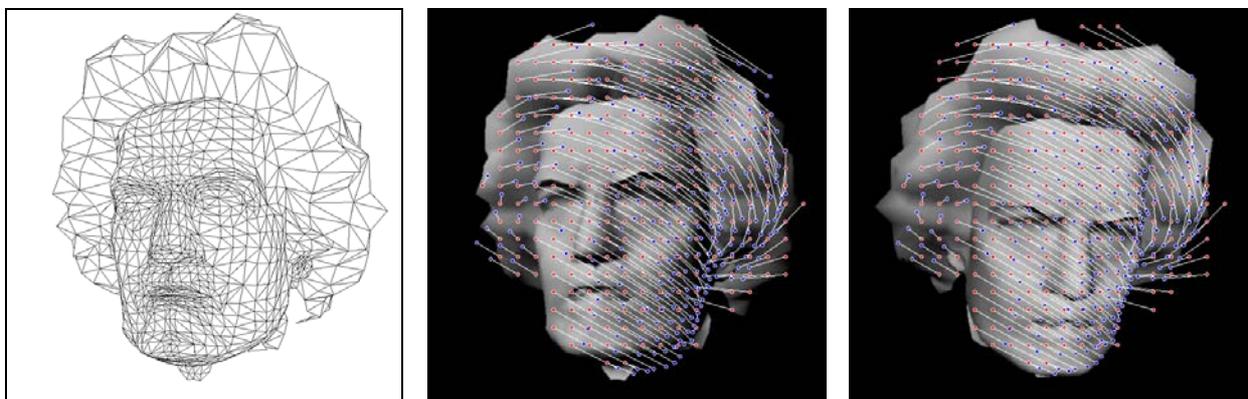
Die digitalisierten Bildausschnitte besitzen eine Auflösung von  $768 \times 1.024$  Pixeln. Aus den jeweils 500 extrahierten Interest-Punkten sind ca. 350 Zuordnungen berechnet worden, was einer Quote von 70 Prozent entspricht. Dabei wurde eine Bild-Pyramide mit drei Stufen verwendet und auf der obersten Ebene die relative Orientierung aus ca. 45 homologen Punkten geschätzt.



**Abb. 3.17:** Ergebnisse der automatischen merkmalsbasierten Korrespondenz-Analyse am Beispiel der Göttin VIKTORIA auf der Siegestsäule in Berlin

Es sollte darauf hingewiesen werden, dass bei der Korrespondenz-Analyse weder Glättungsannahmen noch Längenbeschränkungen verwendet wurden. Die einzige Bedingung ist durch die automatisch berechnete Epipolar-Geometrie definiert.

In Ergänzung zu den realen Bildbeispielen wurde ein Simulator entwickelt [MEH94], der bei einem gegebenen 3D-Modell und einer definierten Bewegung für jeden Punkt die exakte Verschiebung im Bild berechnen kann (vgl. Abb. 3.18). Mit solchen idealen Korrespondenzen lässt sich sowohl die Stabilität als auch die Genauigkeit von Orientierungsverfahren präziser analysieren. Des Weiteren können kritische Bewegungen oder Oberflächen wesentlich unkomplizierter getestet werden.



a) Drahtgitter-Modell

b) Schattiertes linkes Bild

c) Rechtes Stereobild

**Abb. 3.18:** Kontrollierte Korrespondenz-Analyse mit Hilfe von Simulations-Daten am Beispiel einer Drehung des 3D-Modells von BEETHOVEN

## 3.7 Diskussion

Die Korrespondenz-Analyse im Nahbereich stellt ohne Kenntnis der Aufnahmegeometrie eine große Herausforderung dar. Es konnte aber gezeigt werden, dass durch einen merkmalsbasierten Ansatz mit Interest-Punkten und die geeignete Kombination von verschiedenen Korrelations-Verfahren in Verbindung mit einer AuflösungsPyramide eine Lösung gefunden werden kann.

Die Untersuchungen haben gezeigt, dass die geometrisch unabhängigen Merkmale der *differentiellen Invarianten* die Verträglichkeit von farbigen Bildelementen angemessen beschreiben und sich daher gut für eine Vorauswahl eignen. Die nachfolgende Korrespondenz-Analyse wird dadurch stabilisiert und beschleunigt. Durch die *geometrische Anpassung* mit einem affinen Deformationsmodell können Verzerrungen iterativ kompensiert werden. Die vorgeschlagene adaptive Glättung verbessert die Konvergenz und mit der Rotation des Referenzbildes sowie dem symmetrischen Vergleich wird der Wirkungsbereich noch einmal entscheidend vergrößert.

Die für Farbbilder erweiterte *normierte Kreuz-Korrelation* ist sehr robust bei radiometrischen Unterschieden und eignet sich deshalb hervorragend als abschließendes Ähnlichkeitskriterium. Der Erfolg der Korrespondenz-Analyse hängt nun davon ab, ob nach der Merkmalsreduktion über die GAUSSsche *AuflösungsPyramide* mindestens 50 Prozent korrekte Zuordnungen getroffen wurden, so dass die relative Orientierung mit einem robusten Verfahren geschätzt werden kann. Die hierarchische Verfeinerung mit der Definition eines Suchraums erhöht schließlich die Anzahl und Genauigkeit der Zuordnungen.

*„Angesichts von Hindernissen  
mag die kürzeste Linie  
zwischen zwei Punkten  
die krumme sein.“*

BERTOLT BRECHT [1898-1956];  
dt. Schriftsteller und Regisseur, Galilei

## 4 Projektive Stereo-Geometrie

In dem folgenden Kapitel soll die räumliche Lage und Orientierung zwischen zwei Bildern rekonstruiert werden (*Relative Orientierung*). Im Gegensatz zu der in der Photogrammetrie üblichen Verfahrensweise sind die erforderlichen Kamera-Parameter allerdings nicht bekannt. Dieses Problem soll deshalb mit Hilfe der algebraischen *projektiven Geometrie* [SEM52] gelöst werden.

Bei der traditionell verwendeten euklidischen Geometrie bleiben metrische Größen wie z.B. Längenverhältnisse, Winkel, Parallelitäten nach einer euklidischen Transformation (Drehung, Verschiebung oder Skalierung) erhalten. Werden aber zwei parallele Geraden im Raum durch die Perspektiv-Projektion einer Kamera aufgenommen, so treffen sich die abgebildeten Geraden in einem Fluchtpunkt, d.h. bei einer perspektivischen Abbildung bleiben im Allgemeinen weder Parallelitäten, Winkel noch Abstände erhalten. Die projektive Geometrie stellt also eine Erweiterung der euklidischen Geometrie dar und enthält zusätzliche Transformationen zur Beschreibung der Perspektiv-Projektion einer Kamera.

Nach einer kurzen Einführung in die *Grundlagen der projektiven Geometrie* (Abschnitt 4.1) soll zunächst das *geometrische Modell* der verwendeten Kamera erläutert werden (Abschnitt 4.2). Anschließend werden unterschiedliche Methoden zur *relativen Orientierung* eines Bildpaares auf der Grundlage von homologen Bildpunkten besprochen (Abschnitt 4.3). Da die Automatisierung des Verfahrens eine besonders stabile Berechnung erfordert, wird in Abschnitt 4.4 ein neues Verfahren zur robusten Parameterschätzung mit Hilfe von genetischen Algorithmen vorgestellt. Abschließend können die räumlichen Koordinaten der extrahierten Bildpunkte durch Triangulation bestimmt werden (*Räumlicher Vorwärtsschnitt*).

Die Menge der extrahierten homologen Bildpunkte, die beiden aus der relativen Orientierung bestimmten Kamera-Modelle sowie die triangulierten Objektpunkte definieren dann eine so genannte *projektive Rekonstruktion* des ausgewerteten Stereo-Bildpaares.

### 4.1 Grundlagen der projektiven Geometrie

Im Rahmen der vorliegenden Arbeit kann das breite Spektrum der projektiven Geometrie nur ausschnittsweise behandelt werden. Für eine umfassende Darstellung der Thematik sind die Bücher von HARTLEY und ZISSERMANN [HAR00] sowie von FAUGERAS und LUONG [FAU01] zu empfehlen.

Für ein besseres Verständnis der Arbeit soll nach einer kurzen Einführung in verschiedene *geometrische Räume* (Abschnitt 4.1.1) die Repräsentation von Punkten und Linien mit Hilfe von *homogenen Koordinaten* erläutert werden (Abschnitt 4.1.2). Anschließend werden die besonderen Beziehungen zwischen Punkten und Linien in der algebraischen Darstellung anhand von ausgewählten Beispielen veranschaulicht (Abschnitt 4.1.3).

Für die Auto-Kalibrierung in Kapitel 6 sind neben einfachen Punkten und Linien allerdings noch weitere geometrische Repräsentationen notwendig. Daher werden in Abschnitt 4.1.4 die homogenen Darstellungen von *Ebenen*, *Kegelschnitten* und *Flächen zweiter Ordnung* kurz vorgestellt. Abschließend wird noch auf die Berechnung von *projektiven Transformationen* eingegangen, die ein wesentliches Konzept der projektiven Geometrie darstellen (Abschnitt 4.1.5).

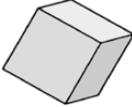
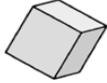
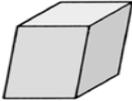
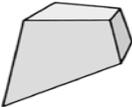
### 4.1.1 Hierarchie geometrischer Räume

Jeder geometrische Raum enthält bestimmte Transformationen und definiert dadurch spezielle Invarianten. Neben dem gebräuchlichen *euklidischen* Raum  $\mathbb{R}^3$  gibt es mit dem *affinen* und dem *metrischen* Raum noch zwei weitere Spezialisierungen des *projektiven* Raums  $\mathbb{P}^3$ . Während sich Objekte im metrischen Raum nur durch einen Skalierungsfaktor vom euklidischen Raum unterscheiden, bleiben z.B. im affinen Raum keine Winkel erhalten. Im projektiven Raum gehen schließlich auch noch Parallelitäten verloren. Je umfassender also die Transformationen für einen geometrischen Raum sind, umso schwächer sind die Aussagen seiner Geometrie (vgl. Tab. 4.1).

Die Hierarchie der geometrischen Transformationen ist in Tab. 4.2 zusammengefasst, wobei **A** eine allgemeine invertierbare  $3 \times 3$ -Matrix mit acht Freiheitsgraden, **R** eine Rotations-Matrix, **t** einen Verschiebungsvektor, **v** einen allgemeinen 3-Vektor und  $\nu$  sowie  $\lambda$  jeweils Skalare bezeichnen. Das dargestellte Beispiel zeigt einen Würfel nach einer geometrischen Transformation mit den entsprechenden Freiheitsgraden (*Degrees of Freedom*, DOF).

Raum	Transformationen				Invarianten			
	Rotation, Translation	Isotrope Skalierung	Skalierung, Scherung	Perspektiv-Projektion	Streckenlängen, Volumen	Längenverhältnisse, Winkel	Teilverhältnisse, Parallelitäten	Inzidenzen, Doppelverhältnisse
<b>Euklidisch</b>	X				X	X	X	X
<b>Metrisch</b>	X	X				X	X	X
<b>Affin</b>	X	X	X				X	X
<b>Projektiv</b>	X	X	X	X				X

Tab. 4.1: Beziehungen zwischen den geometrischen Räumen [FAU01]

Raum	DOF	Transformation	Beispiel
<b>Euklidisch (Kongruenz)</b>	6	$\mathbf{T}_E = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$	
<b>Metrisch (Ähnlichkeit)</b>	7	$\mathbf{T}_M = \begin{bmatrix} \lambda \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$	
<b>Affin</b>	12	$\mathbf{T}_A = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$	
<b>Projektiv</b>	15	$\mathbf{T}_P = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix}$	

Tab. 4.2: Hierarchie räumlicher Transformationen [POL99b]

### 4.1.2 Homogene Koordinaten

Der Einsatz von *homogenen Koordinaten* vereinfacht die Darstellung von räumlichen Beziehungen. Ein wesentlicher Vorteil liegt in der Möglichkeit, mit unendlich fernen (*uneigentlichen* oder *idealen*) Objekten, wie z.B. Punkten, Linien oder Ebenen arbeiten zu können. Die homogenen Koordinaten werden von euklidischen Koordinaten durch Hinzufügen einer zusätzlichen Komponente und einer freien Skalierung abgeleitet.

Die geometrische Repräsentation als Vektor oder Matrix ist also redundant, so dass  $\mathbf{a}$  und  $\lambda \cdot \mathbf{a}$  für  $\lambda \neq 0$  das gleiche Objekt beschreiben. Die homogene Repräsentation von einem *Bildpunkt*  $(x, y)^T$  in der *projektiven Ebene*  $\mathbb{IP}^2$  ist durch den 3-Vektor

$$\mathbf{x} = \lambda \cdot (x, y, 1)^T = (u, v, w)^T \quad (4.1)$$

definiert, wobei  $w$  als *homogene Koordinate* bezeichnet wird und nicht alle Elemente 0 sein dürfen. Die euklidischen Koordinaten können aus den homogenen Koordinaten mit

$$x = \frac{u}{w} \quad \text{und} \quad y = \frac{v}{w} \quad \text{für} \quad w \neq 0 \quad (4.2)$$

zurückgerechnet werden. Ein *idealer Punkt* ist durch  $\mathbf{x}_\infty = (u, v, 0)^T$  gekennzeichnet. Des Weiteren lässt sich ein euklidischer *Objektpunkt*  $(X, Y, Z)^T$  in homogenen Koordinaten mit dem 4-Vektor

$$\mathbf{X} = \lambda \cdot (X, Y, Z, 1)^T = (U, V, W, T)^T \quad (4.3)$$

darstellen und mit

$$X = \frac{U}{T}, \quad Y = \frac{V}{T} \quad \text{und} \quad Z = \frac{W}{T} \quad \text{für} \quad T \neq 0 \quad (4.4)$$

aus dem *projektiven Raum*  $\mathbb{IP}^3$  wieder zurück transformieren. Ein *idealer Objektpunkt* ist durch den Vektor  $\mathbf{X}_\infty = (U, V, W, 0)^T$  definiert. Eine *Gerade* in der *projektiven Ebene*  $\mathbb{IP}^2$  lässt sich auf unterschiedliche Weise definieren, wobei üblicherweise die *HESSEsche Normalform* oder die *Achsenabschnittsform* eingesetzt wird. Bei der *impliziten Form* wird eine Gerade durch die drei Parameter  $a$ ,  $b$  und  $c$  festgelegt:

$$ax + by + c = 0 \quad (4.5)$$

Diese Darstellung kann in homogenen Koordinaten direkt mit dem 3-Vektor

$$\mathbf{l} = \lambda \cdot (a, b, c)^T \quad (4.6)$$

ausgedrückt werden. Die ersten beiden Komponenten von  $\mathbf{l}$  definieren den Normalen-Vektor, so dass die beschriebene Gerade eine Richtung  $(b, -a)^T$  mit dem Winkel  $\theta = \text{atan}(-a/b)$  besitzt. Eine *ideale Gerade* ist in homogenen Koordinaten durch den Vektor  $\mathbf{l}_\infty = (0, 0, 1)^T$  gegeben und stellt anschaulich den Horizont dar.

Die homogene Darstellung von *Geraden* im *projektiven Raum*  $\mathbb{IP}^3$  führt schließlich auf so genannte *PLÜCKER-Koordinaten*, die in einem 6-Vektor zusammengefasst werden. Dieser enthält bei Generierung durch zwei Objektpunkte  $\mathbf{X}$  und  $\mathbf{Y}$  die Differenz  $\mathbf{Y} - \mathbf{X}$  und das Kreuzprodukt  $\mathbf{X} \times \mathbf{Y}$ .

### 4.1.3 Dualität von Punkten und Linien

Die so definierten Punkte und Geraden werden in der Ebene jeweils durch einen 3-Vektor repräsentiert und haben einige interessante Gemeinsamkeiten (*Dualität*). Eine Geradengleichung legt z.B. fest, wann ein Punkt auf einer Geraden liegt oder eine Gerade durch einen Punkt geht, d.h. Punkt und Gerade inzident sind. So liegt in der Ebene ein Punkt  $\mathbf{x}$  genau dann auf der Geraden  $\mathbf{l}$ , wenn die *Inzidenz-Relation*

$$\mathbf{x}^T \mathbf{l} = \mathbf{l}^T \mathbf{x} = ax + by + cw = 0 \quad (4.7)$$

erfüllt ist. Der Schnittpunkt zwischen zwei Geraden ist durch das Kreuzprodukt

$$\mathbf{x} = \mathbf{l}_1 \times \mathbf{l}_2 \quad (4.8)$$

definiert. Eine Linie, welche zwei Punkte verbindet, ist entsprechend durch

$$\mathbf{l} = \mathbf{x}_1 \times \mathbf{x}_2 \quad (4.9)$$

festgelegt. Als Erweiterung liegen drei Punkte auf einer Geraden, wenn die Bedingung

$$\det[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3] = 0 \quad (4.10)$$

erfüllt ist (*Kollinearität*) und schließlich schneiden sich drei Linien in genau einem Punkt, wenn

$$\det[\mathbf{l}_1 \quad \mathbf{l}_2 \quad \mathbf{l}_3] = 0 \quad (4.11)$$

gilt (*Übereinstimmung*).

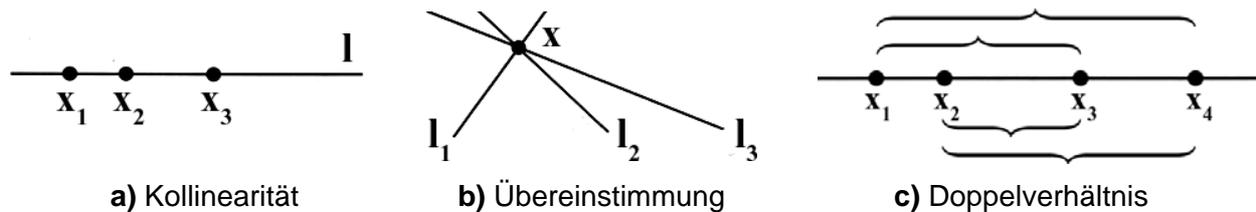


Abb. 4.1: Grundlegende Beziehungen zwischen Punkten und Linien

Eine wichtige Invariante im projektiven Raum ist das *Doppelverhältnis*. Es ist für vier kollineare Punkte folgendermaßen definiert:

$$DV(x_1, x_2, x_3, x_4) = \frac{d(x_1, x_3) \cdot d(x_2, x_4)}{d(x_1, x_4) \cdot d(x_2, x_3)}, \quad (4.12)$$

wobei  $d(x, y)$  den geometrischen Abstand zwischen den Punkten  $x$  und  $y$  bezeichnet. Abb. 4.1 veranschaulicht diese Zusammenhänge.

#### 4.1.4 Geometrische Repräsentationen

In dieser Arbeit werden neben einfachen geometrischen Objekten auch die Grundlagen von *Kegelschnitten* und *Flächen zweiter Ordnung* benötigt. Nach der Erläuterung der homogenen Darstellung und ihrer dualen Repräsentation werden die besonderen Eigenschaften der *idealen Ebene*, des *absoluten Kegelschnitts* und der *dualen absoluten Fläche zweiter Ordnung* dargelegt.

##### 4.1.4.1 Kegelschnitte (Conics)

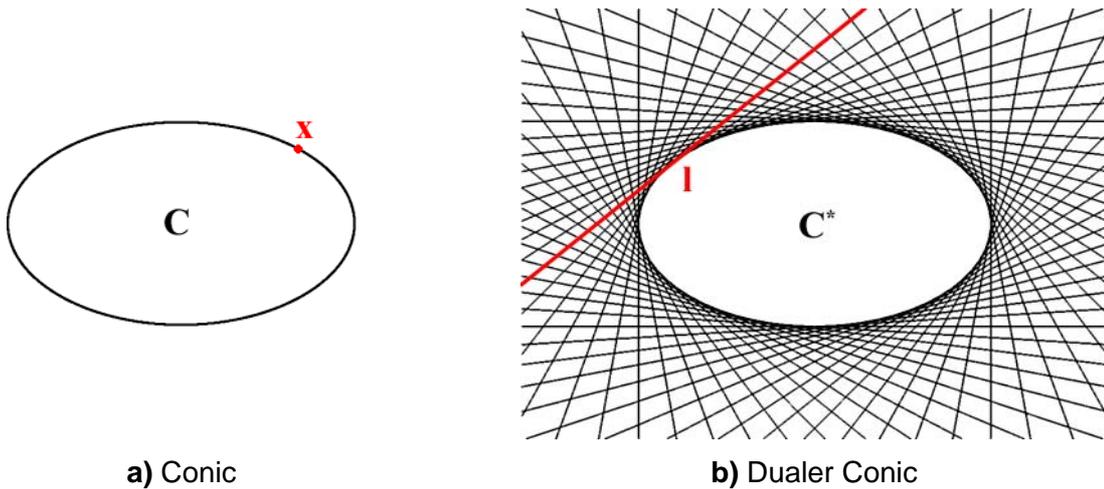
Wenn man eine Kegelfläche mit einer Ebene zum Schnitt bringt, dann ergeben sich so genannte *Kegelschnitte*. Je nach Lage der Ebene entstehen als Schnittkurven z.B. Kreise, Ellipsen, Hyperbeln oder Parabeln. Üblicherweise wird als *Kegelschnitt-Gleichung* das quadratische Polynom

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (4.13)$$

verwendet. In der projektiven Geometrie wird dieser Zusammenhang durch eine homogene  $3 \times 3$ -Matrix repräsentiert. Ein *Conic*  $C$  in der projektiven Ebene  $\mathbb{IP}^2$  wird durch die Menge aller Punkte  $\mathbf{x} = (x, y, w)^T$  definiert, für die

$$\mathbf{x}^T C \mathbf{x} = 0 \quad \text{mit} \quad C = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix} \quad (4.14)$$

gilt. Aufgrund der Dualität zwischen Punkten und Linien (vgl. Abschnitt 4.1.3) kann ein Kegelschnitt auch durch die Menge aller Tangential-Linien  $l$  beschrieben werden (siehe Abb. 4.2).



**Abb. 4.2:** Beispiel für einen Kegelschnitt und die duale Repräsentation

Die *duale Repräsentation* eines Conic  $C^*$  ist dann durch

$$\mathbf{l}^T \mathbf{C}^* \mathbf{l} = 0 \quad (4.15)$$

definiert, wobei die Eigenschaft  $\mathbf{C}^* = \mathbf{C}^{-1}$  gilt, falls  $\mathbf{C}$  nicht singulär ist. Durch die Darstellung in homogenen Koordinaten besitzen die beiden symmetrischen  $3 \times 3$  Matrizen jeweils fünf Freiheitsgrade.

#### 4.1.4.2 Flächen zweiter Ordnung (Quadrics)

Für die Auto-Kalibrierung in Abschnitt 6.4 werden im projektiven Raum  $\mathbb{P}^3$  *Flächen zweiter Ordnung* eingesetzt. Beispiele hierfür sind Kugel, Ellipsoide, Hyperboloide oder Paraboloiden, wie in der Abb. 4.3a dargestellt. Ein *Quadric*  $\mathbf{Q}$  wird durch eine homogene  $4 \times 4$ -Matrix repräsentiert und ist durch die Menge aller Objektpunkte  $\mathbf{X} = (X, Y, Z, W)^T$  definiert, für die

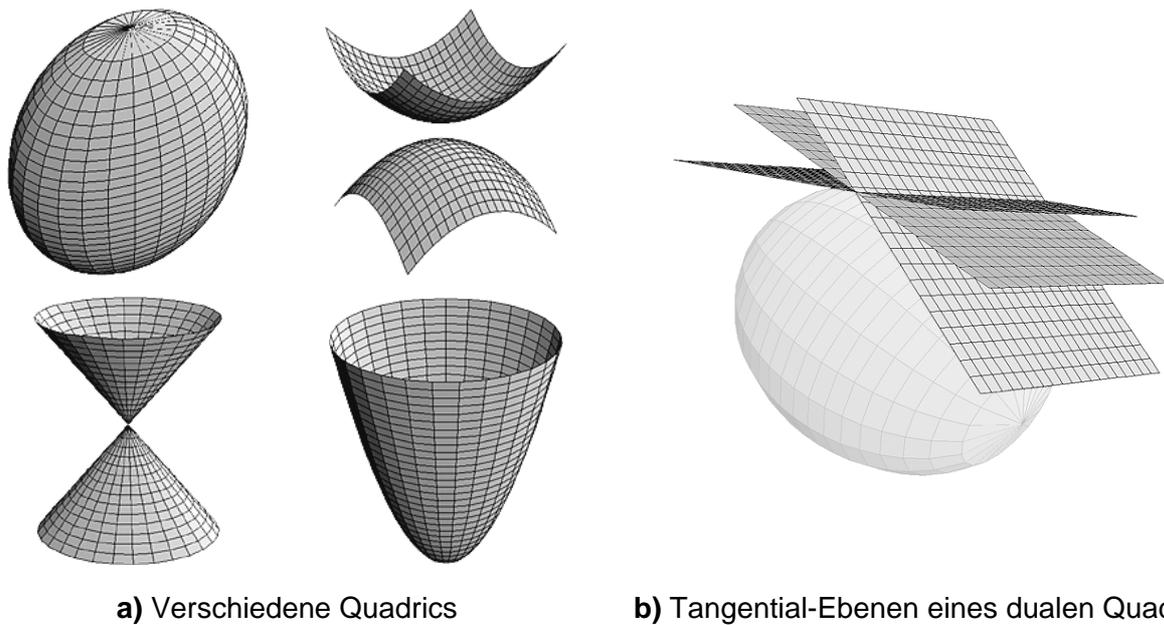
$$\mathbf{X}^T \mathbf{Q} \mathbf{X} = 0 \quad (4.16)$$

gilt. Genau wie bei den Conics gibt es auch bei den Quadrics eine duale Darstellung. So ist ein *dualer Quadric*  $\mathbf{Q}^*$  durch die Menge aller Tangential-Ebenen  $\mathbf{\Pi}$  definiert, für die

$$\mathbf{\Pi}^T \mathbf{Q}^* \mathbf{\Pi} = 0 \quad (4.17)$$

erfüllt ist. Da eine entsprechende Visualisierung nicht ganz trivial ist, werden in Abb. 4.3b nur einige Tangential-Ebenen angedeutet.

Der Zusammenhang zwischen den symmetrischen  $4 \times 4$ -Matrizen  $\mathbf{Q}$  und  $\mathbf{Q}^*$  lautet  $\mathbf{Q}^* = \mathbf{Q}^{-1}$ , falls  $\mathbf{Q}$  nicht singulär ist. Durch die homogene Repräsentation besitzen (duale) Quadrics neun Freiheitsgrade. Ein wesentlicher Vorteil der dualen Repräsentation von Conics und Quadrics ist ihre vereinfachte Transformations-Vorschrift.



**Abb. 4.3:** Beispiele für Flächen zweiter Ordnung und die duale Repräsentation

Die *Transformation*  $\mathbf{T}$  eines Conics oder Quadrics

$$\mathbf{C}' = \mathbf{T}^{-\mathbf{T}} \mathbf{C} \mathbf{T}^{-1} \quad \text{bzw.} \quad \mathbf{Q}' = \mathbf{T}^{-\mathbf{T}} \mathbf{Q} \mathbf{T}^{-1} \quad (4.18)$$

kann aufgrund der notwendigen Matrix-Invertierung schwierig sein, während die Transformation der entsprechend dualen Repräsentation mit

$$\mathbf{C}^{*\prime} = \mathbf{T} \mathbf{C}^* \mathbf{T}^{\mathbf{T}} \quad \text{bzw.} \quad \mathbf{Q}^{*\prime} = \mathbf{T} \mathbf{Q}^* \mathbf{T}^{\mathbf{T}} \quad (4.19)$$

bedeutend einfacher ist.

#### 4.1.4.3 Die ideale Ebene $\Pi_{\infty}$

Eine allgemeine *Ebene im projektiven Raum*  $\mathbb{P}^3$  kann in homogenen Koordinaten durch den 4-Vektor  $\mathbf{\Pi} = (A, B, C, D)^{\mathbf{T}}$  beschrieben werden, wobei die ersten drei Komponenten die Orientierung der Oberflächen-Normale bestimmen. Eine wichtige Funktion bei der affinen Rekonstruktion besitzt die *ideale Ebene* (Plane at Infinity)

$$\mathbf{\Pi}_{\infty} = (0, 0, 0, 1)^{\mathbf{T}}, \quad (4.20)$$

die man sich als unendlich ferne Himmelskugel vorstellen kann (vgl. Abschnitt 6.2).

#### 4.1.4.4 Der absolute Kegelschnitt $\Omega_{\infty}$

Der *absolute Kegelschnitt* (Absolute Conic) definiert mit  $\mathbf{C} = \mathbf{I}$  einen Einheitskreis und liegt in der idealen Ebene  $\mathbf{\Pi}_{\infty}$ . Er besteht nur aus komplexen Punkten und lässt sich mit den beiden Gleichungen

$$\left. \begin{array}{l} X^2 + Y^2 + Z^2 \\ W \end{array} \right\} = 0 \quad (4.21)$$

beschreiben. Obwohl unter einer metrischen Transformation keine punktweise Identität erhalten bleibt, so ist die Form des absoluten Kegelschnitts invariant. Außerdem ist  $\mathbf{Q}_\infty$  in jeder Szene vorhanden und eignet sich daher als virtuelles Kalibrierobjekt für die metrische Rekonstruktion (vgl. Abschnitt 6.2.2).

#### 4.1.4.5 Die duale absolute Fläche zweiter Ordnung $\mathbf{Q}_\infty^*$

Die *duale* Repräsentation der *absoluten Fläche zweiter Ordnung* (Dual Absolute Quadric) ist zunächst eine positiv-definierte symmetrische  $4 \times 4$ -Matrix, die mit zehn Parametern beschrieben werden kann. Da die Matrix mit dem Rang drei singulär ist

$$\det(\mathbf{Q}_\infty^*) = 0, \quad (4.22)$$

handelt es sich um einen degenerierten Quadric. In Verbindung mit der Darstellung in homogenen Koordinaten verbleiben noch acht Freiheitsgrade. Ein wesentlicher Vorteil des dualen absoluten Quadrics  $\mathbf{Q}_\infty^*$  ist, dass er sich dual zum absoluten Conic  $\mathbf{Q}_\infty$  verhält und zusätzlich noch die uneigentliche Ebene  $\mathbf{H}_\infty$  enthält. Für die Auto-Kalibrierung in Abschnitt 6.4 wird die wichtige Eigenschaft ausgenutzt, dass der duale absolute Quadric im metrischen Raum die Form

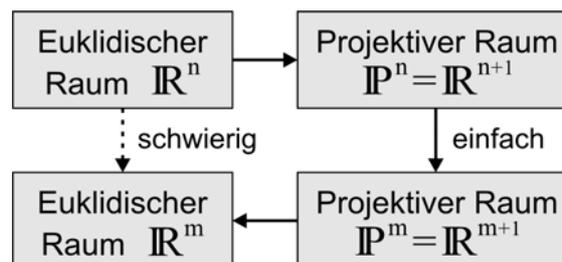
$$\mathbf{Q}_M^* = \text{diag}(1, 1, 1, 0) \quad (4.23)$$

besitzt.

#### 4.1.5 Projektive Transformationen

Ein wesentliches Konzept der projektiven Geometrie liegt in der Möglichkeit, Abbildungen zwischen verschiedenen geometrischen Objekten zu definieren (vgl. Abb. 4.4). Eine allgemeine projektive Transformation, die linear in homogenen Koordinaten und invertierbar ist, wird *Homographie* genannt. Im projektiven Raum  $\mathbb{P}^n$  kann sie durch eine nicht singuläre  $(n+1) \times (n+1)$  homogene Matrix  $\mathbf{H}$  dargestellt werden, die bis auf einen Skalierungsfaktor definiert ist.

Nachfolgend soll die Berechnung von projektiven Transformationen mit Hilfe der *direkten linearen Transformation* (DLT) am Beispiel einer *Homographie in der projektiven Ebene*  $\mathbb{P}^2$  und einer *Homographie im projektiven Raum*  $\mathbb{P}^3$  erläutert werden.



**Abb. 4.4:** Prinzip der Nutzung projektiver Transformationen

#### 4.1.5.1 Direkte lineare Transformation (DLT)

Der DLT-Algorithmus hat seinen Ursprung in der Veröffentlichung von ABDEL-AZIZ und KARARA [ABD71]. Die zugrunde liegende Idee wurde zwar schon 20 Jahre zuvor in der photogrammetrischen Literatur von DAS [DAS49] beschrieben, sie fand aber aufgrund der damals fehlenden Rechenleistung noch wenig Beachtung.

Bei der Kalibrierung von Kameras ist die DLT zwischen Bild- und Raumkoordinatensystem ein häufig eingesetztes Verfahren, wobei sich die Anwendung in der Regel auf den Einbild-Fall mit bekannten Passpunkten eines Kalibrier-Objektes beschränkt. Die typische Bestimmung der elf Transformations-Parameter aus mindestens sechs Passpunkten ist z.B. in KLETTE [KLE96b] ausführlich beschrieben.

Die Möglichkeiten der DLT sollen hier allerdings in einem weitaus allgemeineren Zusammenhang betrachtet werden. Im Vergleich zu iterativen Minimierungsverfahren ist die DLT schneller und benötigt keine initiale Schätzung für Startwerte. Durch die direkte Lösung bestehen keine Konvergenzprobleme und die schwierige Auswahl eines Abbruchkriteriums entfällt ebenfalls. Der DLT-Algorithmus liefert mit einer geeigneten Normierung recht gute Ergebnisse.

Ein Nachteil liegt in der Tatsache, dass ein algebraischer Fehler minimiert wird, der weder geometrisch noch statistisch eine reale Bedeutung hat. Mit der algebraischen Lösung der DLT als Startwert bietet eine Optimierung von geometrischen Abständen in der Regel bessere und anschaulichere Lösungen. Für diese nicht-lineare Optimierung wird häufig das robuste LEVENBERG-MARQUARDT-Verfahren eingesetzt, wobei in dieser Arbeit die mehrdimensionale Minimierung mit dem Ansatz von POWELL [PRE02] durchgeführt wird.

#### 4.1.5.2 Homographie in der projektiven Ebene

Der folgende Abschnitt ist ein Beispiel für die Anwendung der DLT zur Berechnung einer Homographie in der projektiven Ebene  $\mathbb{P}^2$  und stützt sich im Wesentlichen auf die in HARTLEY [HAR00] empfohlene Vorgehensweise. Die  $3 \times 3$ -Homographie-Matrix  $\mathbf{H}$  mit acht Freiheitsgraden soll für  $n \geq 4$  homologe Bildpunkte  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  bestimmt werden, so dass  $\mathbf{x}'_i = \mathbf{H} \mathbf{x}_i$  gilt. Die Punkte dürfen allerdings nicht kollinear verteilt sein, da ansonsten keine eindeutige Lösung existiert.

Nachdem die Bildpunkte in homogene Koordinaten transformiert wurden (vgl. Abschnitt 4.1.2), sollten sie *normiert* werden, da dies zu numerisch stabileren Ergebnissen führt. Die Menge aller Punkte  $\mathbf{x}_i = (x_i, y_i, w_i)^T$  wird zunächst so verschoben, dass der Schwerpunkt der Koordinaten im Ursprung  $\mathbf{o} = (0, 0)^T$  liegt. Anschließend werden sie so skaliert, dass der durchschnittliche Abstand vom Ursprung  $\sqrt{2}$  beträgt.

Diese Normierung kann mit Hilfe einer metrischen Transformation durchgeführt werden (vgl. Tab. 4.2). Beschränkt man sich bei den Parametern auf die Verschiebung  $\mathbf{t} = (t_x, t_y)^T$  sowie die Skalierung  $\boldsymbol{\lambda} = (\lambda_x, \lambda_y)^T$ , dann vereinfacht sich die Berechnung der Punkte zu

$$\hat{\mathbf{x}}_i = \mathbf{T} \mathbf{x}_i \quad \text{mit} \quad \mathbf{T} = \begin{bmatrix} \lambda_x & 0 & t_x \\ 0 & \lambda_y & t_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.24)$$

Experimente in [HAR97a] haben gezeigt, dass auch eine isotrope Skalierung ohne signifikante Einschränkungen möglich ist. Die Normierung muss unabhängig auch für das zweite Bild berechnet werden, um die Punkte  $\mathbf{x}'_i$  mit  $\mathbf{T}'$  nach  $\hat{\mathbf{x}}'_i$  zu transformieren. Für jedes normierte Bildpunkt-Paar  $\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$  müssen nun die beiden Zeilen der Matrix  $\mathbf{A}_i$  aufgestellt

$$\mathbf{A}_i = \begin{bmatrix} \hat{w}'_i \hat{\mathbf{x}}_i^T & \mathbf{0}^T & -\hat{x}'_i \hat{\mathbf{x}}_i^T \\ \mathbf{0}^T & \hat{w}_i \hat{\mathbf{x}}_i^T & -\hat{y}'_i \hat{\mathbf{x}}_i^T \end{bmatrix} \quad (4.25)$$

und anschließend in einer gemeinsamen  $2n \times 9$  großen Design-Matrix  $\mathbf{A}$  zusammengefügt werden. Ab  $n \geq 4$  Korrespondenzen kann das *lineare homogene Gleichungssystem* der Form

$$\mathbf{A} \mathbf{h} = \mathbf{0} \quad \text{mit} \quad \mathbf{h} = (h_1, \dots, h_9)^T \quad (4.26)$$

direkt nach einer Singulärwert-Zerlegung durch die Bestimmung des Rechtseigenvektors von  $\mathbf{A}$ , der dem kleinsten Eigenwert entspricht, gelöst werden (vgl. Anhang A.2.1). Wenn die Komponenten des Lösungsvektors  $\mathbf{h}$  wieder in Matrixform gebracht werden

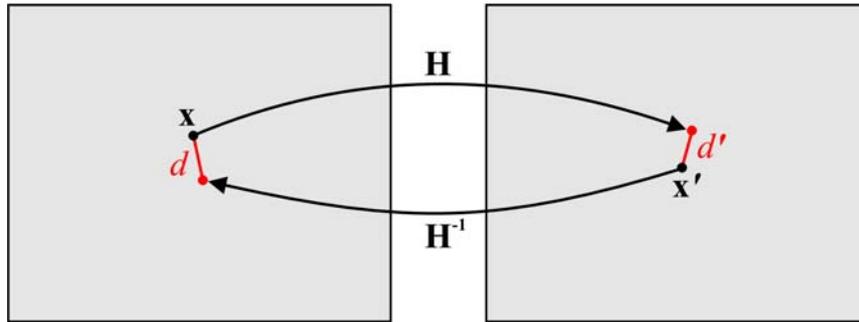
$$\hat{\mathbf{H}} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (4.27)$$

erhält man eine Homographie-Matrix für die normierten Koordinaten. Damit die projektive Transformation auch für die ursprünglichen Koordinaten eingesetzt werden kann, muss zum Abschluss die Normierung mit

$$\mathbf{H} = \mathbf{T}'^{-1} \hat{\mathbf{H}} \mathbf{T} \quad (4.28)$$

wieder rückgängig gemacht werden. Im Prinzip können die zur Normierung und deren Aufhebung verwendeten Transformationen ebenfalls als Homographien angesehen werden.

Mit der algebraischen Lösung als Näherungswert sollte anschließend noch eine *nicht-lineare Optimierung* des geometrischen Fehlers durch Anpassen der neun Matrix-Komponenten durchgeführt werden.



**Abb. 4.5:** Der symmetrische Transfer-Fehler einer Homographie

Für die Homographie in der projektiven Ebene eignet sich der *symmetrische Transfer-Fehler* zwischen beiden Bildern

$$\min_{\mathbf{H}} \sum_i d(\mathbf{x}_i, \mathbf{H}^{-1}\mathbf{x}'_i)^2 + d(\mathbf{x}'_i, \mathbf{H}\mathbf{x}_i)^2, \quad (4.29)$$

wobei  $d(\mathbf{x}, \mathbf{y})$  den geometrischen (euklidischen) Abstand zwischen den Punkten  $\mathbf{x}$  und  $\mathbf{y}$  im Bild bezeichnet (vgl. Abb. 4.5).

### 4.1.5.3 Homographie im projektiven Raum

Genau wie bei der Homographie in der projektiven Ebene kann zum Transfer von Objektpunkten  $\mathbf{X}_i = (X_i, Y_i, Z_i, W_i)^T$  im Raum  $\mathbb{P}^3$  ebenfalls eine projektive Transformation definiert werden. Wenn für  $n \geq 5$  Punkte die euklidischen Raumkoordinaten  $\mathbf{X}'_i = (X'_i, Y'_i, Z'_i, W'_i)^T$  z.B. durch ein Kalibrierobjekt bekannt sind, dann kann eine  $4 \times 4$ -Homographie-Matrix  $\mathbf{H}$  mit 15 Freiheitsgraden bestimmt werden, so dass  $\mathbf{X}'_i = \mathbf{H}\mathbf{X}_i$  gilt. Dadurch ist eine direkte metrische Rekonstruktion möglich (vgl. Abschnitt 6.1).

Handelt es sich bei den Zielkoordinaten  $\mathbf{X}'_i$  allerdings nicht um euklidische, sondern um projektive Koordinaten, so muss die Optimierung von Abständen als grobe Näherung angesehen werden, da in der Regel Abstände im projektiven Raum  $\mathbb{P}^3$  nicht erhalten bleiben (vgl. Abschnitt 4.1.1). Für die Berechnung müssen zunächst die homogenen Koordinaten mit der metrischen Transformation

$$\mathbf{T} = \begin{bmatrix} \lambda_x & 0 & 0 & t_x \\ 0 & \lambda_y & 0 & t_y \\ 0 & 0 & \lambda_z & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.30)$$

in den Ursprung verschoben und auf den mittleren Abstand  $\sqrt{3}$  skaliert werden. Anschließend kann für jeden korrespondierenden Objektpunkt  $\mathbf{X}_i \leftrightarrow \mathbf{X}'_i$  die Matrix

$$\mathbf{A}_i = \begin{bmatrix} W_i \mathbf{X}_i^T & \mathbf{0}^T & \mathbf{0}^T & -X'_i \mathbf{X}_i^T \\ \mathbf{0}^T & W_i \mathbf{X}_i^T & \mathbf{0}^T & -Y'_i \mathbf{X}_i^T \\ \mathbf{0}^T & \mathbf{0}^T & W_i \mathbf{X}_i^T & -Z'_i \mathbf{X}_i^T \\ \hline Y_i \mathbf{X}_i^T & -X'_i \mathbf{X}_i^T & \mathbf{0}^T & \mathbf{0}^T \\ Z'_i \mathbf{X}_i^T & \mathbf{0}^T & -X'_i \mathbf{X}_i^T & \mathbf{0}^T \\ \mathbf{0}^T & Z'_i \mathbf{X}_i^T & -Y'_i \mathbf{X}_i^T & \mathbf{0}^T \end{bmatrix} \quad (4.31)$$

aufgestellt werden. Während die ersten drei Gleichungen die projektive Transformation schon vollständig beschreiben, reduzieren die unteren drei Bedingungen laut CSURKA, DEMIRDJIAN und HORAUD [CSU99] numerische Instabilitäten, die durch eine ungleichmäßige Gewichtung der letzten Spalte entstehen können. Für  $n \geq 5$  Punkte können die einzelnen Matrizen wieder zu einer  $6n \times 16$  großen Design-Matrix  $\mathbf{A}$  zusammengefügt und das lineare homogene Gleichungssystem der Form

$$\mathbf{A}\mathbf{h} = \mathbf{0} \quad \text{mit} \quad \mathbf{h} = (h_1, \dots, h_{16})^T \quad (4.32)$$

direkt mit Hilfe der Singulärwert-Zerlegung gelöst werden (vgl. Anhang A.2.1). Nun müssen die 16 Elemente des Lösungsvektors  $\mathbf{h}$  noch in die  $4 \times 4$ -Matrix-Form gebracht und abschließend die Normierung rückgängig gemacht werden. Zur Verbesserung der algebraischen Lösung kann auch hier eine *nicht-lineare Optimierung* des geometrischen Fehlers durchgeführt werden. Eine geeignete Fehlerfunktion stellt der Abstand  $d$  zwischen den mit  $\mathbf{H}$  transformierten projektiven Punkten und den (euklidischen) Objektpunkten  $\mathbf{X}'_i$  dar. Der *symmetrische Transfer-Fehler* im projektiven Raum kann durch die Optimierung der 16 Komponenten von  $\mathbf{H}$  mit

$$\min_{\mathbf{H}} \sum_i d(\mathbf{X}_i, \mathbf{H}^{-1} \mathbf{X}'_i)^2 + d(\mathbf{X}'_i, \mathbf{H} \mathbf{X}_i)^2 \quad (4.33)$$

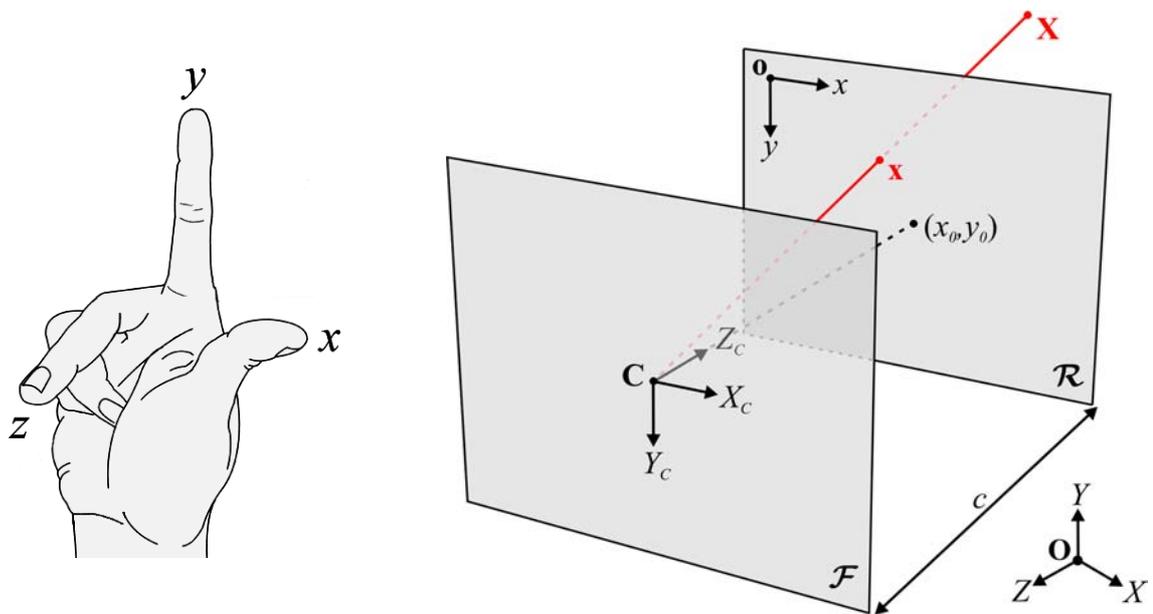
minimiert werden.

## 4.2 Geometrisches Kameramodell

Nach der Einführung in die allgemeinen Grundlagen der projektiven Geometrie soll im Folgenden erläutert werden, wie die Bildentstehung geometrisch modelliert werden kann und welche Abbildungsregeln verwendet werden (Abschnitt 4.2.1).

Um die perspektivische Projektion eines Objektpunktes in die Bildebene beschreiben zu können, müssen zunächst die Parameter des Kameramodells festgelegt werden. In der projektiven Geometrie erfolgt dies üblicherweise mit Hilfe der so genannten *Projektions-Matrix*  $\mathbf{P}$ , die auf dem einfachen Modell einer *Lochkamera* basiert (Abschnitt 4.2.2).

Anschließend werden noch einige interessante Zusammenhänge zwischen der algebraischen Darstellung der Projektions-Matrix und ihrer geometrischen Bedeutung aufgezeigt (Abschnitt 4.2.3).



a) Rechtshändiges Koordinatensystem

b) Perspektivisches Abbildungsmodell

Abb. 4.6: Definition der Koordinatensysteme bei dem verwendeten Lochkamera-Modell

### 4.2.1 Abbildungsmodell

Als mathematisches Abbildungsmodell liegt die kamerazentrierte *Zentralprojektion* zugrunde, wobei sowohl für die Bild- als auch die Raumkoordinaten ein rechtshändiges kartesisches Koordinatensystem angenommen wird.

Das *Projektionszentrum*  $C$  liegt dabei in der *Verschwindungsebene*  $\mathcal{F}$  und die *Projektionsebene*  $\mathcal{R}$  liegt im Abstand der *Kamerakonstanten*  $c$  parallel dazu (siehe Abb. 4.6). Die Zeilen- und Spaltenposition des *Bildhauptpunktes*  $(x_0, y_0)$  wird durch den Schnitt der  $Z_c$ -Achse mit der Projektionsebene definiert und muss nicht zwangsläufig mit dem Bildmittelpunkt identisch sein.

### 4.2.2 Die Projektions-Matrix $\mathbf{P}$

Falls die verwendete Abbildung als projektiv angenommen werden kann, so genügt die Angabe der  $3 \times 4$ -*Projektions-Matrix*  $\mathbf{P}$  zur vollständigen Definition der Kameraeigenschaften. Sie beschreibt die perspektivische Abbildung eines Objektpunktes  $\mathbf{X}$  an die Bildposition  $\mathbf{x}$

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad \text{bzw.} \quad \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} \quad (4.34)$$

und führt somit die Transformation vom projektiven Raum  $\mathbb{IP}^3$  in die Ebene  $\mathbb{IP}^2$  durch.

Aufgrund der Darstellung in homogenen Koordinaten besitzt die Projektions-Matrix  $\mathbf{P}$  nur elf Freiheitsgrade. Ein entscheidender Vorteil dieser kompakten Darstellung in einer Matrix ist, dass die häufige Unklarheit über verwendete Abbildungsregeln, die Reihenfolge von Translationen und Rotationen sowie die Art der Drehwinkel entfällt.

Die Position des Projektionszentrums im Raum sowie die Aufnahme­richtung der Kamera bezüglich eines übergeordneten Welt-Koordinatensystems werden als *äußere Orientierung* bezeichnet. Wenn das *Projektionszentrum*  $\mathbf{C}$  und die *Rotations-Matrix*  $\mathbf{R}$  mit insgesamt sechs Parametern gegeben sind, dann setzt sich  $\mathbf{P}$  folgendermaßen zusammen:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} = \mathbf{K}\mathbf{R}[\mathbf{I} | -\mathbf{C}], \quad (4.35)$$

wobei  $\mathbf{K}$  noch die Parameter der *inneren Orientierung* enthält.

Die  $3 \times 3$ -*Kalibrierungs-Matrix*  $\mathbf{K}$  ist eine obere Dreiecks-Matrix, die aufgrund der homogenen Darstellung fünf Freiheitsgrade besitzt. Ihre Elemente definieren die geometrischen Eigenschaften der Kamera und man bezeichnet eine Kamera als *kalibriert*, wenn die Matrix  $\mathbf{K}$  bekannt ist.

Nachfolgend soll eine kurze Übersicht der gebräuchlichen Kamera-Modelle gegeben werden. Aus technischer Sicht transformiert das *physikalische Modell*

$$\mathbf{K} = \begin{bmatrix} ck_x & -ck_x \cot(\theta) & x_0 \\ 0 & ck_y / \sin(\theta) & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.36)$$

Koordinaten von einem Sensorelement in Zeilen- und Spaltenpositionen des Bildes. Dabei bezeichnen

- $c$  die Kamerakonstante [in Millimetern],
- $k_x$  die Anzahl der Pixel pro Millimeter in Richtung der  $x$ -Achse,
- $k_y$  die Anzahl der Pixel pro Millimeter entlang der  $y$ -Achse,
- $(x_0, y_0)^T$  die Position des Bildhauptpunktes [in Pixeln] und
- $\theta$  den Scherungswinkel zwischen den Bildachsen.

Wenn man bei dem Kamera-Modell von einem konkreten physikalischen Sensor abstrahiert, dann lautet das *algebraische Modell* der Kalibrierungs-Matrix in vereinfachter Schreibweise

$$\mathbf{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4.37)$$

wobei

- $\alpha_x$  der Kamerakonstanten  $c$  in Pixeln entspricht,
- $\gamma = \alpha_y / \alpha_x$  das Seitenverhältnis und
- $s$  als Scherungsfaktor die Rechtwinkligkeit der  $x$ - und  $y$ -Achse

definieren. Im Gegensatz zu diesem vereinfachten geradentreuen Abbildungsmodell arbeiten Kameras üblicherweise mit Objektiven, die Sichtstrahlen bündeln, um eine höhere Beleuchtungsstärke zu erreichen. Der Nachteil bei der Verwendung von Objektiven sind optische Abbildungsfehler (*Verzeichnung*), die die geometrische Ähnlichkeit des Bildes verfälschen.

Neben der Objektivverzeichnung, die häufig als symmetrische radiale Abweichung modelliert werden kann, sind auch weitere systematische Fehler beim Abtasten von Videosignalen oder auch zufällige Fehler durch Unebenheit des Films möglich [WIL94, MEH94]. Die Modellierung von solchen nicht-linearen Bildfehler-Funktionen  $\delta_x$  und  $\delta_y$  sind nicht in dem Modell der Projektions-Matrix  $\mathbf{P}$  enthalten, können aber damit kombiniert werden. Die häufig verwendeten algebraischen Kamera-Modelle sind in Tab. 4.3 zusammengefasst.

Kamera-Typ	Kalibrierungs-Matrix
Normiert	$\mathbf{K}_N = \mathbf{I}$
Ideal	$\mathbf{K}_I = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Euklidisch	$\mathbf{K}_E = \begin{bmatrix} \alpha & 0 & x_0 \\ 0 & \alpha & y_0 \\ 0 & 0 & 1 \end{bmatrix}$
Projektiv	$\mathbf{K}_P = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$
Allgemein	$\mathbf{K}_A = \begin{bmatrix} \alpha_x & s & x_0 + \delta_x(x, y) \\ 0 & \alpha_y & y_0 + \delta_y(x, y) \\ 0 & 0 & 1 \end{bmatrix}$

Tab. 4.3: Beispiele gebräuchlicher Kamera-Modelle

Kamera-Typ	Orientierungs-Parameter							DOF
	Äußere					Innere		
	<b>C</b>	<b>R</b>	$\alpha$	$x_0, y_0$	$\gamma$	$s$	$\delta_x, \delta_y$	
<b>Normiert</b>	-	<b>I</b>	1	0, 0	1	0	0, 0	3
<b>Ideal</b>	-	-	-					7
<b>Euklidisch</b>	-	-	-	-	-	-	-	9
<b>Projektiv</b>	-	-	-	-	-	-	-	11
<b>Allgemein</b>	-	-	-	-	-	-	-	>11

**Tab. 4.4:** Annahmen über Orientierungs-Parameter bei den verschiedenen Kamera-Modellen

Eine Übersicht der verschiedenen Modell-Annahmen und der noch verbleibenden Freiheitsgrade (DOF) ist in Tab. 4.4 dargestellt. So wird z.B. die Projektions-Matrix einer projektiven Kamera durch insgesamt elf Parameter definiert, wobei sechs die äußere und fünf die innere Orientierung beschreiben. Die Diskussion, welches Kamera-Modell für das Ziel der Auto-Kalibrierung zweckmäßig ist, wird in Abschnitt 6.1 fortgeführt.

### 4.2.3 Interpretation von $\mathbf{P}$

Zunächst soll die kompakte algebraische Darstellung der Projektions-Matrix durch eine geometrische Deutung veranschaulicht werden. Anschließend werden relativ einfache Methoden zur Extraktion der Orientierungs-Parameter vorgestellt. So wird vorgeschlagen, das Projektionszentrum mit Hilfe einer Singulärwert-Zerlegung und die Kalibrierungs- bzw. Rotations-Matrizen durch eine RQ-Zerlegung zu berechnen.

Da die Projektions-Matrix  $\mathbf{P}$  durch die homogene Darstellung allerdings nur bis auf einen Skalierungsfaktor  $\lambda$  bekannt ist, sollte sie generell vor einer Interpretation normiert werden. Um den Betrag und das Vorzeichen des *Normierungsfaktors* zu bestimmen, muss die erste  $3 \times 3$ -Teilmatrix  $\mathbf{M}$  von  $\mathbf{P} = [\mathbf{M} | \mathbf{t}]$  betrachtet werden. Sei  $\mathbf{m}^3$  die dritte Zeile von  $\mathbf{M}$ , so muss zunächst die gesamte Projektions-Matrix  $\mathbf{P}$  durch den Betrag dieses 3-Vektors geteilt werden, damit die Eigenschaft

$$\|\mathbf{m}^3\| = 1 \quad (4.38)$$

erfüllt ist. Das korrekte Vorzeichen erhält man schließlich durch die Bedingung

$$\det(\mathbf{M}) > 0. \quad (4.39)$$

Bei einer negativen Determinante muss das Vorzeichen aller Komponenten von  $\mathbf{P}$  invertiert werden.

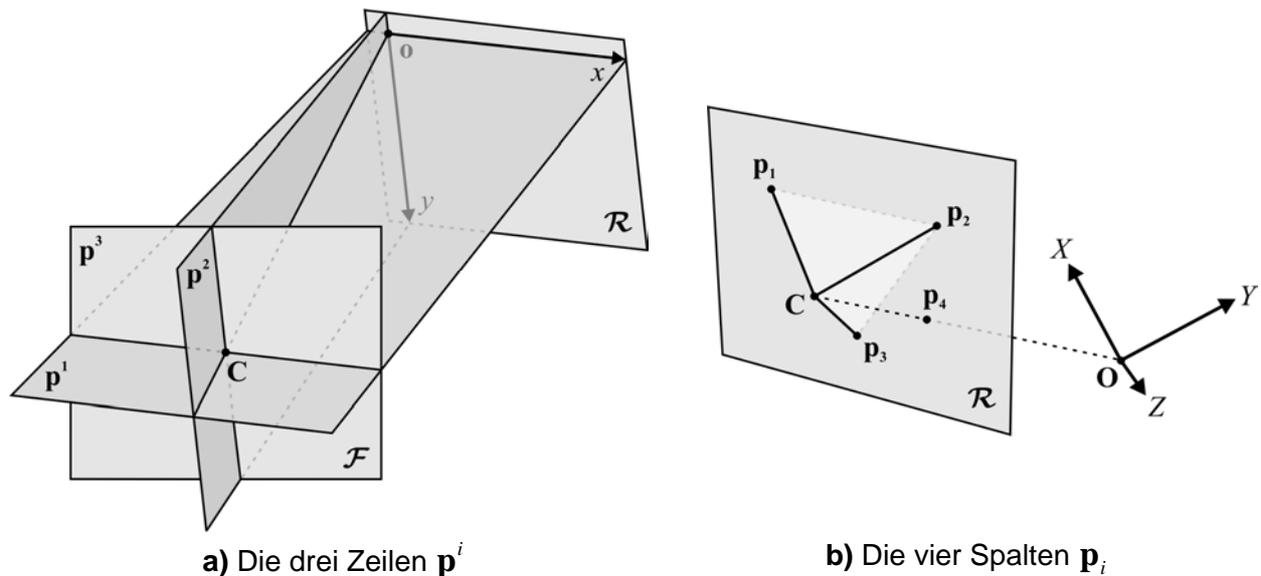


Abb. 4.7: Geometrische Deutung von Zeilen und Spalten der Projektions-Matrix  $\mathbf{P}$

#### 4.2.3.1 Geometrische Deutung

Die Zeilen und Spalten der algebraischen Projektions-Matrix  $\mathbf{P}$  haben eine besondere Bedeutung. So sind die *Zeilen*  $\mathbf{p}^i$  4-Vektoren und können als *Ebenen* im projektiven Raum  $\mathbb{P}^3$  angesehen werden (vgl. Abschnitt 4.1.4.3). Die drei Ebenen schneiden sich im Projektionszentrum  $\mathbf{C}$  und die dritte Zeile  $\mathbf{p}^3$  entspricht der Verschwindungsebene  $\mathcal{F}$ . Die Lage der zu den Zeilen  $\mathbf{p}^1$  und  $\mathbf{p}^2$  gehörenden Ebenen hängt je nach Wahl des Bildkoordinatensystems von der  $x$ - bzw.  $y$ -Achse ab (siehe Abb. 4.7a). Die Schnittgerade dieser beiden Ebenen verbindet das Projektionszentrum mit dem Bildursprung  $\mathbf{o}$ .

Die *Spalten*  $\mathbf{p}_i$  sind 3-Vektoren und stellen besondere *Punkte* in der Projektionsebene  $\mathcal{R}$  dar. Die ersten drei Spalten  $\mathbf{p}_1$ ,  $\mathbf{p}_2$  und  $\mathbf{p}_3$  sind die Abbildungen der Achsenrichtungen des Weltkoordinatensystems und entsprechen daher den Fluchtpunkten der  $X$ -,  $Y$ - bzw.  $Z$ -Achse (vgl. Abb. 4.7b). Die letzte Spalte  $\mathbf{p}_4$  ist die Abbildung des Ursprungs vom Weltkoordinatensystem  $\mathbf{O}$ .

#### 4.2.3.2 Das Projektionszentrum $\mathbf{C}$

Wenn die Projektions-Matrix  $\mathbf{P}$  bekannt ist, so ist das Projektionszentrum  $\mathbf{C} = (c_1, c_2, c_3, c_4)^T$  algebraisch durch

$$c_i = (-1)^i \det(\hat{\mathbf{P}}_i) \quad \text{für } i = 1, \dots, 4 \quad (4.40)$$

definiert, wobei  $\hat{\mathbf{P}}_i$  die Matrix  $\mathbf{P}$  ohne die  $i$ -te Spalte bezeichnet.

Das Projektionszentrum  $\mathbf{C}$  ist allerdings auch der besondere Punkt, für den  $\mathbf{P}\mathbf{C} = \mathbf{0}$  gilt. Wenn man diese Eigenschaft als lineares homogenes Gleichungssystem auffasst, dann kann das Projektionszentrum viel eleganter und numerisch stabiler mit Hilfe der Singu-

lärwert-Zerlegung von  $\mathbf{P}$  bestimmt werden (vgl. Anhang A.2.1). In der Praxis ist zu beachten, dass die Rechtecks-Matrix  $\mathbf{P}$  für die SVD um eine Zeile mit Nullen ergänzt werden muss.

### 4.2.3.3 Parameter-Extraktion durch RQ-Zerlegung

Um aus einer gegebenen Projektions-Matrix  $\mathbf{P} = [\mathbf{M} \mid \mathbf{t}]$  die Kalibrierungs-Matrix  $\mathbf{K}$  und die Rotations-Matrix  $\mathbf{R}$  zu extrahieren, muss lediglich die erste  $3 \times 3$ -Teilmatrix  $\mathbf{M}$  betrachtet werden. In der Literatur existieren dazu viele unterschiedliche Methoden, wobei die *RQ-Zerlegung* eine besonders elegante Variante darstellt [HAR00].

Das RQ-Theorem besagt, dass eine nicht-singuläre Matrix als Produkt einer oberen *Dreiecks-Matrix*  $\mathbf{R}$  und einer *Orthogonal-Matrix*  $\mathbf{Q}$  faktorisiert werden kann, d.h.

$$\mathbf{M} = \mathbf{R}\mathbf{Q} \quad \text{bzw.} \quad \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}. \quad (4.41)$$

Es muss betont werden, dass in diesem Abschnitt aufgrund der üblichen mathematischen Konvention die Orthogonal-Matrix  $\mathbf{Q}$  die *Rotations-Matrix* und die obere Dreiecks-Matrix  $\mathbf{R}$  die *Kalibrierungs-Matrix* bezeichnen.

Die Berechnung der RQ-Zerlegung kann z.B. mit den GIVENS-Rotationen durchgeführt werden. Dabei setzt die Multiplikation von  $\mathbf{M}$  mit der Matrix

$$\mathbf{Q}_X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}, \quad s = -\frac{m_{32}}{t} \quad \text{und} \quad c = \frac{m_{33}}{t} \quad \text{mit} \quad t = \sqrt{m_{32}^2 + m_{33}^2} \quad (4.42)$$

das Element  $m_{32}$  auf 0 und durch die anschließende Multiplikation mit

$$\mathbf{Q}_Y = \begin{bmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{bmatrix}, \quad s = \frac{m_{31}}{t} \quad \text{und} \quad c = \frac{m_{33}}{t} \quad \text{mit} \quad t = \sqrt{m_{31}^2 + m_{33}^2} \quad (4.43)$$

von rechts wird auch  $m_{31}$  zu 0. Die zweite Spalte von  $\mathbf{M}$  wird bei dieser Operation nicht verändert. Die Multiplikation mit

$$\mathbf{Q}_Z = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad s = -\frac{m_{21}}{t} \quad \text{und} \quad c = \frac{m_{22}}{t} \quad \text{mit} \quad t = \sqrt{m_{21}^2 + m_{22}^2} \quad (4.44)$$

setzt schließlich auch  $m_{21}$  auf 0. Die ersten beiden Spalten werden durch Linearkombinationen ersetzt, so dass die Elemente  $m_{31}$  und  $m_{32}$  unverändert bleiben. Durch dieses Vorgehen ist aus  $\mathbf{M}$  die obere Dreiecks-Matrix  $\mathbf{R}$  entstanden.

Die Kombination der dabei verwendeten Rotations-Matrizen ergibt schließlich die gesuchte Orthogonal-Matrix  $\mathbf{Q}$

$$\mathbf{R} = \mathbf{M} \mathbf{Q}_X \mathbf{Q}_Y \mathbf{Q}_Z \quad \text{und} \quad \mathbf{Q} = \mathbf{Q}_Z^T \mathbf{Q}_Y^T \mathbf{Q}_X^T. \quad (4.45)$$

Um eine eindeutige Zerlegung zu erhalten, muss die Form der Dreiecks-Matrix  $\mathbf{R}$  allerdings noch eingeschränkt werden. Unter der Bedingung, dass die Diagonalelemente positiv sein müssen, kann die Mehrdeutigkeit bei der Zerlegung aber aufgehoben werden. Um ein negatives Diagonalelement  $r_{ii}$  der Dreiecks-Matrix anzupassen, müssen lediglich die Vorzeichen der gesamten Zeile in  $\mathbf{R}$  sowie der entsprechenden Spalte in  $\mathbf{Q}$  invertiert werden.

Die bei der Parametrisierung der drei GIVENS-Rotationen eingesetzten Parameter  $s = \sin(\theta)$  und  $c = \cos(\theta)$  definieren jeweils einen Drehwinkel  $\theta$ . Die einzelnen Winkel  $\theta_x$  (*roll*),  $\theta_y$  (*pitch*) und  $\theta_z$  (*yaw*) entsprechen den EULER-Winkeln  $(\omega, \varphi, \kappa)$  in der XYZ-Notation.

#### 4.2.3.4 Numerisches Beispiel

Um die Extraktion der Orientierungs-Parameter zu veranschaulichen, soll ein numerisches Beispiel gegeben werden. Die exemplarische Projektions-Matrix

$$\mathbf{P} = \left[ \begin{array}{ccc|c} -559,425 & -69,134 & 25,301 & -998,081 \\ -68,914 & -418,097 & 355,103 & 388,586 \\ -0,451 & 0,211 & 0,866 & -0,679 \end{array} \right] \quad (4.46)$$

besitzt nach der Berechnung des linearen homogenen Gleichungssystems das Projektionszentrum  $\mathbf{C} = (-1,912, 0,881, -0,427, 1,0)^T$ . Die Kalibrierungs- und Rotations-Matrizen sehen nach der RQ-Zerlegung mit den GIVENS-Rotationen folgendermaßen aus

$$\mathbf{M} = \mathbf{K} \mathbf{R} = \left[ \begin{array}{ccc} 500,19 & 22,84 & 260,12 \\ 0,00 & 492,80 & 250,61 \\ 0,00 & 0,00 & 1,00 \end{array} \right] \left[ \begin{array}{ccc} -0,887 & -0,204 & -0,412 \\ 0,089 & -0,955 & 0,279 \\ -0,451 & 0,211 & 0,866 \end{array} \right], \quad (4.47)$$

wobei das Seitenverhältnis 0,985 beträgt.

## 4.3 Orientierung eines Bildpaares

Der folgende Abschnitt befasst sich mit der gegenseitigen Orientierung von Bildpaaren (*Relative Orientierung*). Falls die relative Lage und Richtung der Bildkoordinatensysteme zueinander bekannt ist, so lässt sich für einen Bildpunkt die Position in dem anderen Bild wesentlich einschränken. Dies wird durch Ausnutzen der *Epipolar-Geometrie* erreicht, die nur von den Kamera-Parametern und nicht von dem abgebildeten Objekt abhängt (Abschnitt 4.3.1).

In der projektiven Geometrie wird die Epipolar-Geometrie durch die so genannte *Fundamental-Matrix* repräsentiert, die sich aus wenigen homologen Bildpunkten bestimmen lässt (Abschnitt 4.3.2). Neben der einfachen Berechnung mit dem *normierten 8-Punkt Algorithmus* (Abschnitt 4.3.3) und dem schnellen *minimalen 7-Punkt Algorithmus* (Abschnitt 4.3.4) soll auch die Verbesserung durch eine *nicht-lineare Optimierung* beschrieben werden (Abschnitt 4.3.5).

Ist durch die Bestimmung der Fundamental-Matrix die relative Orientierung der Bilder erst einmal bekannt, dann können im unkalibrierten Fall zwei Projektions-Matrizen abgeleitet werden, welche die zugrunde liegende Abbildungsgeometrie projektiv beschreiben. Abschließend wird noch kurz auf die besonderen Eigenschaften der *essentiellen Matrix* eingegangen, die eine Spezialisierung der Fundamental-Matrix darstellt (Abschnitt 4.3.7).

### 4.3.1 Epipolar-Geometrie

Wenn ein Objektpunkt  $\mathbf{X}$  gleichzeitig in zwei Bildern abgebildet wird, dann gibt es einige grundlegende Beziehungen, die mit Hilfe der *Epipolar-Geometrie* beschrieben werden können (vgl. Abb. 4.8). Die Gerade, die beide Projektionszentren  $\mathbf{C}$  und  $\mathbf{C}'$  miteinander verbindet, wird *Basislinie* genannt. Ihr Schnittpunkt mit einer Projektionsebene, der einer Abbildung des jeweils anderen Projektionszentrums entspricht, wird als *Epipol*  $\mathbf{e}$  bezeichnet.

Um für einen Bildpunkt  $\mathbf{x}$  die geometrische Position im zweiten Bild zu bestimmen, lässt sich die Suche auf eine Gerade reduzieren, die so genannte *Epipolar-Gerade*  $\mathbf{l}'$  für  $\mathbf{x}$ . Alle möglichen Epipolar-Geraden dieses Bildes treffen sich in dem Epipol  $\mathbf{e}'$ . Umgekehrt gilt diese geometrische Beziehung natürlich auch für  $\mathbf{x}'$ ,  $\mathbf{l}$  und  $\mathbf{e}$ . In der projektiven Geometrie kann dieser Zusammenhang ohne Kenntnis von Kamera-Parametern mit Hilfe der projektiven Fundamental-Matrix  $\mathbf{F}$  genutzt werden.

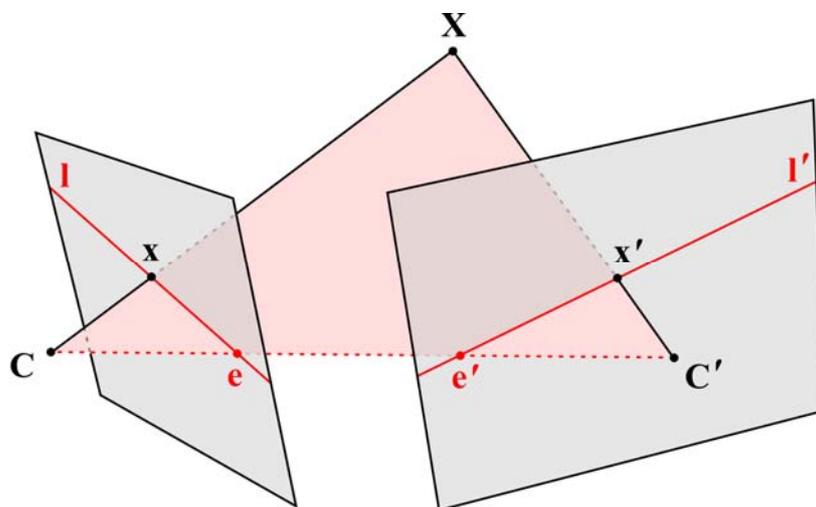


Abb. 4.8: Epipolar-Geometrie für ein Bildpaar

### 4.3.2 Die Fundamental-Matrix $\mathbf{F}$

Die *Fundamental-Matrix*  $\mathbf{F}$  (Bifokal-Tensor) enthält die gesamte Information der projektiven relativen Orientierung und ist deren algebraische Repräsentation. Sie bezieht sich direkt auf die ursprünglichen gemessenen Zeilen- und Spaltenpositionen und erfordert keine vorige Transformation in spezielle Bildkoordinaten. Die Fundamental-Matrix ist zunächst eine homogene  $3 \times 3$ -Matrix, die mit dem Rang zwei singulär ist. Sie besitzt zwar neun Elemente, aber nur sieben Freiheitsgrade und ist damit überparametrisiert.

Da  $\mathbf{F}$  durch homogene Koordinaten nur bis auf einen Skalierungsfaktor bestimmt ist, sind zunächst nur acht Parameter relevant. Die Elemente der Fundamental-Matrix sind allerdings nicht unabhängig, sondern durch die *Singularitäts-Bedingung*  $\det(\mathbf{F}) = 0$  verknüpft. Diese wichtige Beziehung gewährleistet, dass sich alle Epipolar-Geraden genau in einem Punkt schneiden. Von einem allgemeinen Standpunkt betrachtet ist  $\mathbf{F}$  invariant unter einer projektiven Transformation, die 16 Komponenten besitzt (vgl. Tab. 4.2). Jede Projektions-Matrix  $\mathbf{P}$  enthält 12 Komponenten und da aufgrund der homogenen Koordinaten die globale Skalierung irrelevant ist, ergeben sich für die projektiven Beziehungen zwischen zwei Bildern insgesamt  $2 \times 11 - 15 = 7$  Freiheitsgrade.

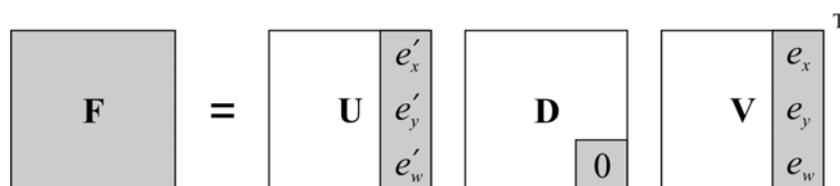
Die Fundamental-Matrix  $\mathbf{F}$  definiert für einen Bildpunkt  $\mathbf{x}$  in dem anderen Bild eine entsprechende Gerade  $\mathbf{l}'$  als Abbildung des Projektionsstrahls und ermöglicht dadurch eine *Prädiktion* des Punktes  $\mathbf{x}'$ . Die beiden *Epipolar-Geraden* sind durch

$$\mathbf{l}' = \mathbf{F}\mathbf{x} \quad \text{bzw.} \quad \mathbf{l} = \mathbf{F}^T\mathbf{x}' \quad (4.48)$$

bestimmt. Für diese projektive Transformation  $\mathbf{x} \mapsto \mathbf{l}'$  gibt es allerdings keine inverse Abbildung  $\mathbf{l}' \mapsto \mathbf{x}$  und das erklärt, warum die Matrix  $\mathbf{F}$  singulär ist. Die beiden *Epipole*  $\mathbf{e}$  und  $\mathbf{e}'$  sind spezielle Punkte in der Projektionsebene, an denen die Epipolar-Geraden undefiniert sind

$$\mathbf{F}\mathbf{e} = \mathbf{0} \quad \text{und} \quad \mathbf{F}^T\mathbf{e}' = \mathbf{0}. \quad (4.49)$$

Diese Eigenschaft kann als lineares homogenes Gleichungssystem aufgefasst und die Epipole daher mit der Singulärwert-Zerlegung aus der Fundamental-Matrix abgeleitet werden (vgl. Anhang A.2.1). Um beide Epipole zu bestimmen, ist allerdings nur eine Zerlegung notwendig, da sich neben  $\mathbf{e}$  als rechtem Nullvektor der zweite Epipol  $\mathbf{e}'$  sofort durch den linken Nullvektor von  $\mathbf{F}$  ergibt (siehe Abb. 4.9).



**Abb. 4.9:** Extraktion der beiden Epipole  $\mathbf{e}$  und  $\mathbf{e}'$  durch die Singulärwert-Zerlegung der Fundamental-Matrix  $\mathbf{F}$

Für alle Punkte, die auf den Epipolar-Geraden liegen, gilt die Inzidenz-Relation  $\mathbf{x}^T \mathbf{I} = \mathbf{x}'^T \mathbf{I}' = 0$ . Daraus folgt mit der Gleichung (4.48), dass auch die folgende bilineare *Komplanaritäts-Bedingung* erfüllt sein muss

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0. \quad (4.50)$$

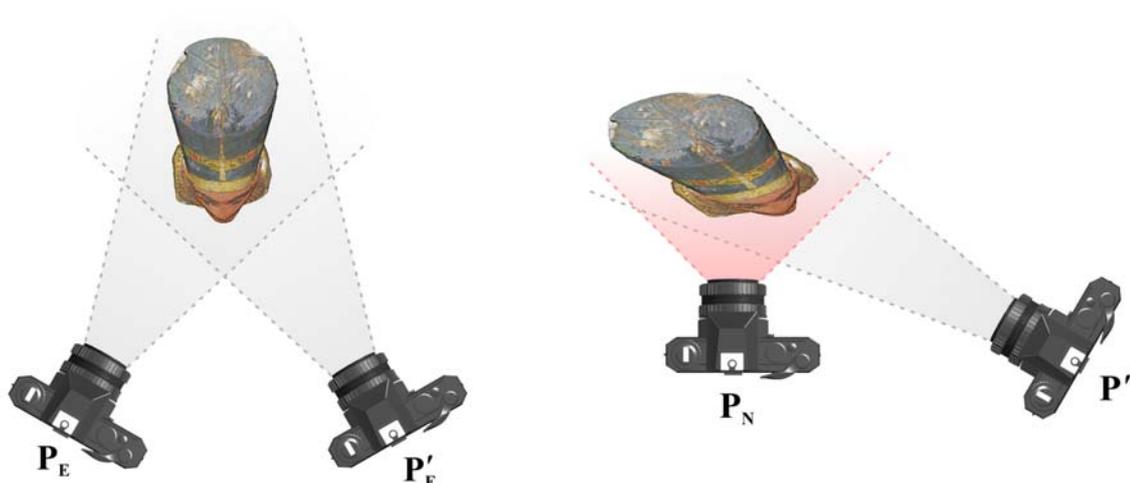
In der Praxis sind die gemessenen Positionen der Bildpunkte jedoch durch kleine Messfehler gestört. Des Weiteren werden die Epipolar-Geraden durch Transformation mit der Fundamental-Matrix ermittelt, die in der Regel auch nicht aus fehlerfreien homologen Punkten berechnet wurde.

Aufgrund der *Unsicherheit* bei der Lokalisierung eines Bildpunktes und der Berechnung der Fundamental-Matrix sollte der angesprochene Suchbereich oberhalb und unterhalb der Epipolar-Geraden vergrößert werden. Das so genannte *Epipolar-Band* kann je nach Vertrauen in die Position der Epipolar-Geraden unterschiedlich breit sein. CSURKA u.a. [CSU97] haben ein recht aufwändiges analytisches Verfahren zur Berechnung der Unsicherheit der Fundamental-Matrix und ihrer Epipole beschrieben, wobei das resultierende Epipolar-Band auf beiden Seiten durch eine Hyperbel begrenzt wird. In der Regel kann das Epipolar-Band aufgrund der sehr geringen Krümmung auch durch zwei Geraden approximiert werden.

Die Bestimmung der Fundamental-Matrix  $\mathbf{F}$  und der damit verbundenen relativen Orientierung basiert allein auf einer Menge von homologen Bildpunkten, ohne die Orientierungs-Parameter der Kameras explizit zu kennen. Für die numerische Berechnung werden in Abschnitt 4.3.3, 4.3.4 und 4.3.5 noch einige Verfahren vorgestellt. Wenn  $\mathbf{F}$  aber erst einmal bekannt ist, dann lassen sich zwei Projektions-Matrizen definieren

$$\mathbf{P} = [\mathbf{I} | \mathbf{0}] \quad \text{und} \quad \mathbf{P}' = [[\mathbf{e}']_{\times} \mathbf{F} | \mathbf{e}'], \quad (4.51)$$

die die zugrunde liegende Abbildungsgeometrie projektiv beschreiben (vgl. Abb. 4.10).



a) Aufnahme-Konfiguration

b) Die normierte und eine projektive Kamera

**Abb. 4.10:** Beispiel für die projektive Rekonstruktion der Aufnahmegeometrie

Dabei bezeichnet

$$[\mathbf{v}]_{\times} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \quad (4.52)$$

das Kreuzprodukt eines 3-Vektors  $\mathbf{v}$  mit einer  $3 \times 3$ -Matrix anhand der *schiefssymmetrischen Matrix*.

Die erste Projektions-Matrix  $\mathbf{P}$  kann ohne Einschränkung der Allgemeinheit als normierte Kamera angenommen werden (siehe Tab. 4.4) und die zweite Projektions-Matrix  $\mathbf{P}'$  ist nur von der Fundamental-Matrix abhängig, da der Epipol  $\mathbf{e}'$  auch aus  $\mathbf{F}$  bestimmt werden kann (vgl. Gleichung 4.49).

### 4.3.3 Normierter 8-Punkt-Algorithmus

Zunächst soll eine recht einfache, aber dennoch effektive Berechnung der Fundamental-Matrix  $\mathbf{F}$  dargestellt werden. Das Ziel besteht darin, für  $n \geq 8$  homologe Bildpunkte  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  eine Matrix  $\mathbf{F}$  zu finden, so dass die Epipolar-Bedingung  $\mathbf{x}'_i{}^T \mathbf{F} \mathbf{x}_i = 0$  erfüllt ist. Die Idee für den so genannten *8-Punkt-Algorithmus* stammt von LONGUET-HIGGINS [LON81].

Genau wie bei der Bestimmung einer Homographie müssen zunächst die Bildpunkte  $\mathbf{x}$  und  $\mathbf{x}'$  mit den metrischen Transformationen  $\mathbf{T}$  bzw.  $\mathbf{T}'$  normiert werden (vgl. Abschnitt 4.1.5.2). Anschließend kann für die *normierten* homogenen Koordinaten  $\hat{\mathbf{x}}$  und  $\hat{\mathbf{x}}'$  das lineare homogene Gleichungssystem  $\mathbf{A} \mathbf{f} = \mathbf{0}$  mit

$$\mathbf{A} = \begin{bmatrix} \hat{x}_1 \hat{x}'_1 & \hat{y}_1 \hat{x}'_1 & \hat{x}'_1 & \hat{x}_1 \hat{y}'_1 & \hat{y}_1 \hat{y}'_1 & \hat{y}'_1 & \hat{x}_1 & \hat{y}_1 & 1 \\ \vdots & \vdots \\ \hat{x}_n \hat{x}'_n & \hat{y}_n \hat{x}'_n & \hat{x}'_n & \hat{x}_n \hat{y}'_n & \hat{y}_n \hat{y}'_n & \hat{y}'_n & \hat{x}_n & \hat{y}_n & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_9 \end{bmatrix} \quad (4.53)$$

aufgestellt werden. Die Berechnung der Kleinste-Quadrate-Lösung mit Hilfe der DLT (siehe Anhang A.2.1) liefert die Fundamental-Matrix für die normierten Punkte

$$\hat{\mathbf{F}} = \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix}. \quad (4.54)$$

Aufgrund von Messfehlern führt das Lösen des linearen Gleichungssystems im Allgemeinen aber nicht dazu, dass die Matrix  $\hat{\mathbf{F}}$  wie gefordert singulär wird. Wenn  $\hat{\mathbf{F}}$  aber den vollen Rang besitzt, dann schneiden sich die Epipolar-Geraden nicht in einem Punkt. Um dieses Problem zu korrigieren, kann  $\hat{\mathbf{F}}$  durch eine Matrix  $\hat{\mathbf{F}}'$  ersetzt werden, die eben diese Singularitäts-Bedingung  $\det(\hat{\mathbf{F}}') = 0$  erfüllt und möglichst ähnlich ist, d.h.  $\|\hat{\mathbf{F}} - \hat{\mathbf{F}}'\|$  minimiert.

TSAI und HUANG [TSA84] schlugen hierfür eine elegante Methode unter Verwendung der Singulärwert-Zerlegung vor. Nach einer SVD von  $\hat{\mathbf{F}} = \mathbf{U}\mathbf{D}\mathbf{V}^T$  ergibt sich die verbesserte Matrix  $\hat{\mathbf{F}}'$  durch Eliminieren des kleinsten Elementes in der Diagonal-Matrix  $\mathbf{D}$  und die anschließende Rekombination

$$\hat{\mathbf{F}}' = \mathbf{U} \operatorname{diag}(d_1, d_2, 0) \mathbf{V}^T. \quad (4.55)$$

Untersuchungen von PONCE und GENC [PON98] lassen allerdings vermuten, dass das Forcieren der Singularitäts-Bedingung mit einem 2-Stufen-Verfahren noch leicht verbessert werden kann.

Zum Schluss wird die gesuchte Fundamental-Matrix  $\mathbf{F}$  für die Originaldaten  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  durch Aufhebung der Normierung von  $\hat{\mathbf{F}}'$  bestimmt

$$\mathbf{F} = \mathbf{T}'^T \hat{\mathbf{F}}' \mathbf{T}. \quad (4.56)$$

#### 4.3.4 Minimaler 7-Punkt-Algorithmus

Um die Fundamental-Matrix  $\mathbf{F}$  bestimmen zu können, sind eigentlich nur sieben homologe Punkte  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  erforderlich. Das Erfüllen der Singularitäts-Bedingung ist allerdings quadratisch in den Parametern von  $\mathbf{F}$ , so dass letztendlich bis zu drei Lösungen untersucht werden müssen (siehe Gleichung 4.59). Der Vorteil liegt aber im geringeren Berechnungsaufwand im Vergleich zum 8-Punkt-Algorithmus, denn bei der Verwendung von Monte-Carlo-Methoden (vgl. Anhang A.3.3) steigt die Anzahl der Versuche exponentiell mit der Größe des minimalen Datensatzes [HAR00].

Die Lösung der Gleichung (4.53) mit nur  $n = 7$  Punkten führt zu einer unterbestimmten  $7 \times 9$  Design-Matrix  $\mathbf{A}$ . Eine Lösung lässt sich aber dennoch aus dem *zweidimensionalen Nullraum* von  $\mathbf{A}$  berechnen. Seien  $\mathbf{u}$  bzw.  $\mathbf{v}$  die rechten Nullvektoren von  $\mathbf{A}$ , die jeweils dem Singulärwert 0 entsprechen und  $\mathbf{U}$  bzw.  $\mathbf{V}$  die  $3 \times 3$ -Matrix-Repräsentationen dieser beiden Vektoren. Dann ist die gesuchte Fundamental-Matrix  $\mathbf{F}$  durch den *Lösungsraum*

$$\mathbf{F}_t = t\mathbf{U} + (1-t)\mathbf{V} \quad (4.57)$$

definiert, wobei  $t$  einen unbekanntem Skalar bezeichnet. Mit Hilfe der *Singularitäts-Bedingung*

$$\det(\mathbf{F}) = \det(t\mathbf{U} + (1-t)\mathbf{V}) = 0 \quad (4.58)$$

lässt sich dieser unendliche Lösungsraum wieder einschränken. Durch Ausmultiplizieren der  $3 \times 3$ -Determinante erhält man das folgende kubische Polynom in  $t$

$$p(t) = at^3 + bt^2 + ct + d = 0 \quad (4.59)$$

mit den Koeffizienten

$$\begin{aligned}
a &= V_{11}(V_{23}V_{32}-V_{22}V_{33})+V_{12}(V_{21}V_{33}-V_{23}V_{31})+V_{13}(V_{22}V_{31}-V_{21}V_{32}) \\
&+ U_{11}(U_{22}U_{33}-U_{23}U_{32}+U_{23}V_{32}+U_{32}V_{23}-U_{22}V_{33}-U_{33}V_{22}+V_{22}V_{33}-V_{23}V_{32}) \\
&+ U_{12}(U_{23}U_{31}-U_{21}U_{33}+U_{21}V_{33}+U_{33}V_{21}-U_{23}V_{31}-U_{31}V_{23}+V_{23}V_{31}-V_{21}V_{33}) \\
&+ U_{13}(U_{21}U_{32}-U_{22}U_{31}+U_{22}V_{31}+U_{31}V_{22}-U_{21}V_{32}-U_{32}V_{21}+V_{21}V_{32}-V_{22}V_{31}) \\
&+ U_{21}(U_{33}V_{12}-U_{32}V_{13}+V_{13}V_{32}-V_{12}V_{33})+U_{31}(V_{12}V_{23}-V_{13}V_{22}) \\
&+ U_{22}(U_{31}V_{13}-U_{33}V_{11}+V_{11}V_{33}-V_{13}V_{31})+U_{32}(V_{13}V_{21}-V_{11}V_{23}) \\
&+ U_{23}(U_{32}V_{11}-U_{31}V_{12}+V_{12}V_{31}-V_{11}V_{32})+U_{33}(V_{11}V_{22}-V_{12}V_{21}) \\
b &= 3(V_{11}(V_{22}V_{33}-V_{23}V_{32})+V_{12}(V_{23}V_{31}-V_{21}V_{33})+V_{13}(V_{21}V_{32}-V_{22}V_{31})) \\
&+ 2(U_{31}(V_{13}V_{22}-V_{12}V_{23})+U_{32}(V_{11}V_{23}-V_{13}V_{21})+U_{33}(V_{12}V_{21}-V_{11}V_{22})) \\
&+ U_{11}(2(V_{23}V_{32}-V_{22}V_{33})+U_{22}V_{33}+U_{33}V_{22}-U_{23}V_{32}-U_{32}V_{23}) \\
&+ U_{12}(2(V_{21}V_{33}-V_{23}V_{31})+U_{23}V_{31}+U_{31}V_{23}-U_{21}V_{33}-U_{33}V_{21}) \\
&+ U_{13}(2(V_{22}V_{31}-V_{21}V_{32})+U_{21}V_{32}+U_{32}V_{21}-U_{22}V_{31}-U_{31}V_{22}) \\
&+ U_{21}(2(V_{12}V_{33}-V_{13}V_{32})+U_{32}V_{13}-U_{33}V_{12}) \\
&+ U_{22}(2(V_{13}V_{31}-V_{11}V_{33})+U_{33}V_{11}-U_{31}V_{13}) \\
&+ U_{23}(2(V_{11}V_{32}-V_{12}V_{31})+U_{31}V_{12}-U_{32}V_{11}) \\
c &= 3(V_{11}(V_{23}V_{32}-V_{22}V_{33})+V_{12}(V_{21}V_{33}-V_{23}V_{31})+V_{13}(V_{22}V_{31}-V_{21}V_{32})) \\
&+ U_{11}(V_{22}V_{33}-V_{23}V_{32})+U_{12}(V_{23}V_{31}-V_{21}V_{33})+U_{13}(V_{21}V_{32}-V_{22}V_{31}) \\
&+ U_{21}(V_{13}V_{32}-V_{12}V_{33})+U_{22}(V_{11}V_{33}-V_{13}V_{31})+U_{23}(V_{12}V_{31}-V_{11}V_{32}) \\
&+ U_{31}(V_{12}V_{23}-V_{13}V_{22})+U_{32}(V_{13}V_{21}-V_{11}V_{23})+U_{33}(V_{11}V_{22}-V_{12}V_{21}) \\
d &= V_{11}(V_{22}V_{33}-V_{23}V_{32})+V_{12}(V_{23}V_{31}-V_{21}V_{33})+V_{13}(V_{21}V_{32}-V_{22}V_{31}).
\end{aligned} \tag{4.60}$$

Durch die numerische Berechnung der Nullstellen des Polynoms (z.B. mit ZRHQR [PRE02]) erhält man entweder eine oder drei reelle Lösungen. Nach dem Einsetzen der Ergebnisse in Gleichung (4.57) muss aus den möglichen drei Lösungen schließlich diejenige Fundamental-Matrix mit dem geringsten *geometrischen Fehler* ausgewählt werden (siehe Gleichung 4.61).

### 4.3.5 Nicht-lineare Optimierung von $\mathbf{F}$

Ist durch eines der linearen Verfahren erst einmal eine initiale Schätzung für die Fundamental-Matrix  $\mathbf{F}$  vorhanden, so sollte anschließend anstelle des algebraischen besser ein *geometrischer Fehler* betrachtet werden.

Ein gutes Fehlermaß für die Bestimmung von  $\mathbf{F}$  ist der von ZHANG [ZHA98] vorgeschlagene *symmetrische Epipolar-Abstand* (vgl. Abb. 4.11). Dabei werden in beiden Bildern die senkrechten Abstände der Bildpunkte von ihrer entsprechenden Epipolar-Geraden ermittelt

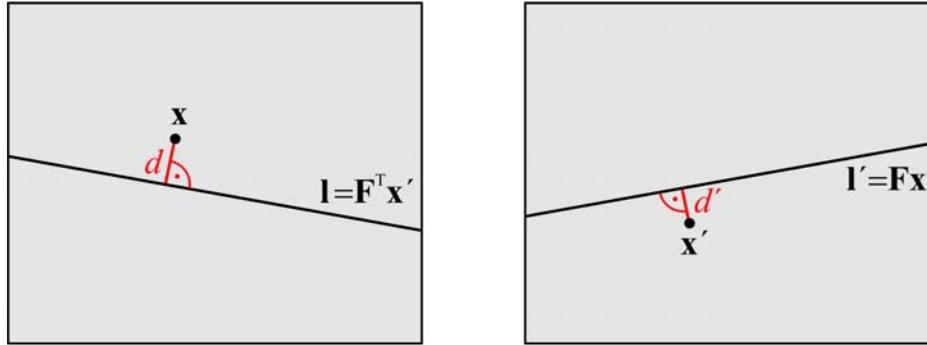


Abb. 4.11: Symmetrischer Epipolar-Abstand

$$\sum_i d(\mathbf{x}'_i, \mathbf{F} \mathbf{x}_i)^2 + d(\mathbf{x}_i, \mathbf{F}^T \mathbf{x}'_i)^2 = \sum_i (\mathbf{x}'_i{}^T \mathbf{F} \mathbf{x}_i)^2 \left( \frac{1}{(\mathbf{F} \mathbf{x}_i)_1^2 + (\mathbf{F} \mathbf{x}_i)_2^2} + \frac{1}{(\mathbf{F}^T \mathbf{x}'_i)_1^2 + (\mathbf{F}^T \mathbf{x}'_i)_2^2} \right), \quad (4.61)$$

wobei  $(\mathbf{F} \mathbf{x}_i)_j^2$  das Quadrat der  $j$ -ten Komponente des Vektors  $\mathbf{l}'_i = \mathbf{F} \mathbf{x}_i$  bezeichnet.

Diese Formulierung des geometrischen (euklidischen) Fehlers ist nahezu identisch mit der häufig verwendeten *Sampson-Approximation* erster Ordnung [SAM82]. Bei der Implementierung muss aber darauf geachtet werden, dass die Fehlerfunktion für die beiden Epipole nicht definiert ist, da in diesem Fall der Nenner 0 wird.

Die nicht-lineare Optimierung des geometrischen Fehlers durch direktes Anpassen der Matrix-Komponenten ist bei der Fundamental-Matrix allerdings nicht so unproblematisch wie bei den Homographien in Abschnitt 4.1.5. Die Elemente sind weder unabhängig noch besitzen sie den gleichen Einfluss und zusätzlich muss die Singularitäts-Bedingung gewährleistet sein.

Eine bessere Variante besteht in der *indirekten* Minimierung des geometrischen Fehlers durch Anpassen der zweiten Projektions-Matrix  $\mathbf{P}'$  sowie der Menge aller Objektpunkte  $\mathbf{X}_i$  [HAR00]. Als Fehlerfunktion eignet sich der geometrische (euklidische) Abstand  $d$  der *gemessenen* homologen Bildpunkte  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  zu den *idealen* Bildpunkten  $\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$ , welche die Bedingung  $\hat{\mathbf{x}}_i{}^T \mathbf{F} \hat{\mathbf{x}}_i = 0$  exakt erfüllen

$$\min_{\mathbf{P}', \mathbf{X}_i} \sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2. \quad (4.62)$$

Für die Initialisierung müssen aber zunächst die Variablen  $\hat{\mathbf{x}}_i$  und  $\hat{\mathbf{x}}'_i$  bestimmt werden. Ausgehend von der geschätzten Fundamental-Matrix  $\mathbf{F}$  können die beiden Projektions-Matrizen  $\mathbf{P}$  und  $\mathbf{P}'$  definiert werden (vgl. Gleichung 4.51). Danach werden aus den gemessenen homologen Bildpunkten  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  durch *Triangulation* die Objektpunkte  $\mathbf{X}_i$  bestimmt (vgl. Abschnitt 4.5).

Die zu  $\mathbf{F}$  konsistenten idealen Bildpunkte erhält man schließlich durch die Projektion dieser Objektpunkte in beide Bilder

$$\hat{\mathbf{x}}_i = \mathbf{P} \mathbf{X}_i \quad \text{und} \quad \hat{\mathbf{x}}'_i = \mathbf{P}' \mathbf{X}_i. \quad (4.63)$$

Nach dieser Initialisierung kann der geometrische Abstand durch nicht-lineare Optimierung der Elemente von  $\mathbf{P}'$  und  $\mathbf{X}_i$  iterativ minimiert werden. Um die numerische Stabilität der Berechnung zu verbessern, sollten die Elemente der  $n$  projektiv bestimmten Objektpunkte  $\mathbf{X}_i$  vom Betrag angeglichen werden. Da sich in der Regel die  $Z$ -Komponente von den übrigen Elementen am meisten unterscheidet, empfiehlt sich eine Normierung auf die dritte Komponente  $\mathbf{X}_i = (X_i, Y_i, 1, W_i)^T$ .

Zusammen mit den zwölf Komponenten der Projektions-Matrix  $\mathbf{P}'$  werden für die Parametrisierung also insgesamt  $3n + 12$  Elemente verwendet

$$(P'_{11}, \dots, P'_{34}, X_1, Y_1, W_1, \dots, X_n, Y_n, W_n)^T. \quad (4.64)$$

Da diese Vorgehensweise einen relativ hohen Berechnungsaufwand erfordert, wird von FAUGERAS [FAU01] eine sequentielle Optimierung

$$\min_{\mathbf{P}'} \left\{ \min_{\mathbf{X}_i} \sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2 \right\} \quad (4.65)$$

und von HARTLEY [HAR00] eine beschleunigte Variante des LEVENBERG-MARQUARDT-Verfahrens vorgeschlagen, die speziell für schwach besetzte Matrizen angepasst ist. Nach der Optimierung kann die gesuchte Fundamental-Matrix aus der verbesserten Projektions-Matrix  $\mathbf{P}' = [\mathbf{M} | \mathbf{t}]$  mit Hilfe der schiefsymmetrischen Matrix (vgl. Gleichung 4.52) folgendermaßen bestimmt werden

$$\mathbf{F} = [\mathbf{t}]_x \mathbf{M}. \quad (4.66)$$

Obwohl  $\mathbf{P}'$  mit ihren zwölf Komponenten eigentlich für die Berechnung der sieben Parameter von  $\mathbf{F}$  überbestimmt ist, gewährleistet die singuläre Matrix  $[\mathbf{t}]_x$  den Rang zwei der Fundamental-Matrix. Unter der Annahme einer GAUSSschen Verteilung der Messfehler ist diese Optimierung des geometrischen Bildfehlers statistisch optimal und entspricht einer minimalen photogrammetrischen Bündel-Ausgleichung (vgl. Abschnitt 5.3).

### 4.3.6 Degenerierte Konfigurationen

Im Prinzip ist die Epipolar-Geometrie immer eindeutig definiert, sofern die beiden Projektionszentren  $\mathbf{C}$  und  $\mathbf{C}'$  nicht übereinstimmen. Bei der Wahl der  $n$  homologen Bildpunkte  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  sind aber trotz einer robusten Parameterschätzung numerisch instabile Situationen möglich. Solche degenerierten Konfigurationen bei der Bestimmung der Fundamental-Matrix können anhand der Dimension des Nullraumes  $\mathbf{N}$  klassifiziert werden (vgl. Tab. 4.5).

$N$	Lösungen	Konfiguration der homologen Bildpunkte
1	Eindeutig	<ul style="list-style-type: none"> <li>Bei <math>n \geq 8</math> perfekten Korrespondenzen in allgemeiner Lage</li> </ul>
2	Eine oder Drei	<ul style="list-style-type: none"> <li>Für <math>n = 7</math> perfekte Korrespondenzen in allgemeiner Lage</li> <li>Wenn alle <math>n &gt; 7</math> Objektpunkte sowie das Projektionszentrum <math>\mathbf{C}</math> auf einer <i>kritischen Oberfläche</i> [MAY93] liegen</li> </ul>
3	2-Parameter-Familie	<p>Wenn die <math>n \geq 6</math> perfekten Korrespondenzen durch eine Homographie <math>\mathbf{x}'_i = \mathbf{H} \mathbf{x}_i</math> beschreibbar sind, z.B.</p> <ul style="list-style-type: none"> <li>Bei einer Rotation um das Projektionszentrum <math>\mathbf{C}</math> (<i>Degenerierte Bewegung</i>)</li> <li>Wenn alle Objektpunkte auf einer Ebene <math>\mathbf{\Pi}</math> liegen (<i>Degenerierte Struktur</i>)</li> </ul>

Tab. 4.5: Degenerierte Konfigurationen bei der Bestimmung von  $\mathbf{F}$  [HAR00]

### 4.3.7 Die essentielle Matrix $\mathbf{E}$

Die *essentielle Matrix*  $\mathbf{E}$  ist eine Spezialisierung der Fundamental-Matrix  $\mathbf{F}$  mit weniger Freiheitsgraden und zusätzlichen Eigenschaften. Die grundlegende Idee der essentiellen Matrix wurde schon 1908 bei VON SANDEN [SAN08] verwendet. Dabei handelt es sich um eine kalibrierte Version von  $\mathbf{F}$ , wobei der Einfluss der bekannten Kalibrierungs-Matrix  $\mathbf{K}$  von der Kamera eliminiert wird

$$\hat{\mathbf{P}} = \mathbf{R} [\mathbf{I} \mid -\mathbf{C}]. \quad (4.67)$$

Die Projektion mit der Matrix  $\hat{\mathbf{P}}$  ist dann für die *reduzierten Bildpunkte*

$$\hat{\mathbf{x}} = \mathbf{K}^{-1} \mathbf{x} \quad (4.68)$$

definiert. Die essentielle Matrix ist genau wie  $\mathbf{F}$  eine homogene  $3 \times 3$  Matrix mit dem Rang zwei, so dass auch die Singularitäts-Bedingung  $\det(\mathbf{E}) = 0$  gilt. Sie besitzt allerdings nur fünf Freiheitsgrade und realisiert die bilineare Komplanaritäts-Bedingung für die reduzierten Bildpunkte

$$\hat{\mathbf{x}}'^T \mathbf{E} \hat{\mathbf{x}} = 0. \quad (4.69)$$

Eine sehr nützliche Eigenschaft von  $\mathbf{E}$  liegt in der Tatsache, dass bei einer Singulärwert-Zerlegung die ersten beiden Singulärwerte identisch sind

$$\mathbf{E} = \mathbf{U} \operatorname{diag}(d, d, 0) \mathbf{V}^T. \quad (4.70)$$

Diese Bedingung wird bei der Auto-Kalibrierung in Abschnitt 6.3.2 verwendet. Die Berechnung der essentiellen Matrix kann bei einer gegebenen Fundamental-Matrix  $\mathbf{F}$  mit

$$\mathbf{E} = \mathbf{K}'^T \mathbf{F} \mathbf{K} \quad (4.71)$$

relativ einfach durchgeführt werden, wobei  $\mathbf{K}$  und  $\mathbf{K}'$  die Kalibrierungs-Matrizen mit der inneren Orientierung der beiden bekannten Kameras bezeichnen.

## 4.4 Robuste Parameterschätzung

Bei der Bestimmung einer projektiven Transformation, wie z.B. einer Homographie  $\mathbf{H}$  in der projektiven Ebene  $\mathbb{P}^2$  bzw. dem projektiven Raum  $\mathbb{P}^3$  oder auch der Fundamental-Matrix  $\mathbf{F}$  sind die verwendeten Messwerte in der Regel nicht fehlerfrei. Um die Berechnung robust zu gestalten, müssen zwei verschiedene Fehlerquellen unterschieden werden:

- **Kleine Fehler (*Rauschen*)**

Die Position der Bildpunkte kann selbst mit einer interaktiven Kontrolle nicht perfekt gemessen werden. Für die so verrauschten Messdaten kann aber im Allgemeinen eine GAUSSsche Fehlerverteilung angenommen und daher eine *nicht-lineare Optimierung* eingesetzt werden (vgl. Abschnitt 4.1.5.2, 4.1.5.3 und 4.3.5).

- **Grobe Fehler (*Ausreißer*)**

Ein gravierendes Problem, das besonders bei automatischen Messungen auftritt, sind grobe Fehler durch falsche Korrespondenzen. Für eine robuste Schätzung der Parameter wird häufig der Fehlereinfluss begrenzt (vgl. *M-Schätzer* im Anhang A.3.1) oder durch *zufälliges Probieren* eine Minimallösung gesucht, die durch möglichst viele Messwerte bestätigt wird (vgl. *RANSAC-Verfahren* im Anhang A.3.3). Diese Vorgehensweise ist zwar relativ unempfindlich gegenüber Ausreißern, jedoch auch sehr rechenintensiv.

Im Folgenden soll ein neuer Ansatz vorgestellt werden, der durch weitgehend *systematisches Probieren* die robuste Parameterschätzung erheblich beschleunigt. Nach einer kurzen Einführung, wie die Berechnung von projektiven Transformationen durch einen *genetischen Algorithmus* verbessert werden kann (Abschnitt 4.4.1), folgt anschließend das beschleunigte Verfahren zur robusten Bestimmung der Fundamental-Matrix (Abschnitt 4.4.2). Obwohl bei dieser Methode grobe Zuordnungsfehler beseitigt und kleine Messungenauigkeiten ausgeglichen werden, sind numerische Instabilitäten nicht ganz auszuschließen. Solche *degenerierten Konfigurationen* sind im Abschnitt 4.4.3 dargestellt.

### 4.4.1 Genetischer Algorithmus GASAC

Ein biologisch motivierter Ansatz zur Lösung von Optimierungproblemen ist der *genetische Algorithmus*, der den erfolgreichen Funktionsprinzipien der Evolution nachempfunden ist. Bereits 1973 hat RECHENBERG [REC73] mit der *Mutations-Selektions-Strategie* ein einfaches Verfahren zur evolutionären Optimierung entwickelt.

Die Grundidee des genetischen Algorithmus stammt von HOLLAND [HOL75] und basiert darauf, die Parameter für ein Problem als Bauplan eines Lebewesens (*Chromosom*) zu betrachten.

Unter den vorgegebenen Umweltbedingungen entscheidet die Überlebensfähigkeit (*Fitness*) in Kombination mit evolutionären Veränderungen am Bauplan über die Zusammensetzung einer besser angepassten Generation. Für den vorliegenden Anwendungsfall wird der Algorithmus in Anlehnung an das RANSAC-Verfahren mit GASAC (*Genetic Algorithm SAMpling Consensus*) bezeichnet.

#### 4.4.1.1 Repräsentation des Genpools

Eine Population  $\mathbf{G}$  besteht aus mehreren Individuen, die jeweils durch ein einzelnes Chromosom charakterisiert werden. Ein solches *Chromosom*  $\mathbf{g} = (g_1, \dots, g_m)$  entspricht einem Tupel von  $m$  Elementen, die als *Gene* bezeichnet werden. Die Repräsentation des vorliegenden Problems durch ein Chromosom ist trivial, wenn man bei  $n$  homologen Punkten  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  jeweils den Index  $i$  direkt als Gen definiert

$$g_k \in \{1, \dots, n\} \quad \text{für } k = 1, \dots, m, \quad (4.72)$$

wobei innerhalb des Chromosoms ein Gen nur genau einmal vorkommen darf.

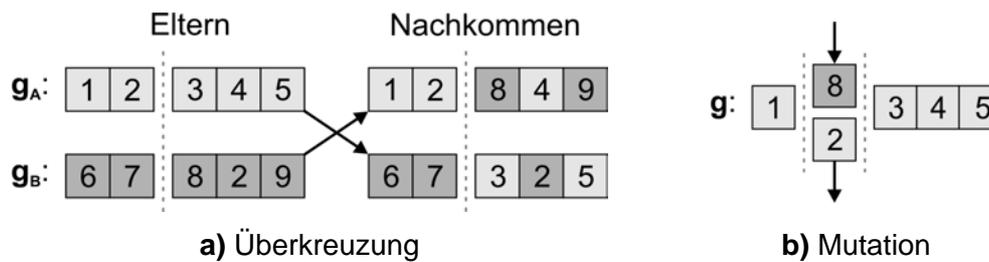
Bei dieser minimalen Parametrisierung bleibt die Länge eines Chromosoms konstant. Für die Homographie  $\mathbf{H}$  in der projektiven Ebene  $\mathbb{IP}^2$  ist  $m = 4$  und in dem projektiven Raum  $\mathbb{IP}^3$  ist  $m = 5$ . Bei dem minimalen 7-Punkt-Algorithmus zur Bestimmung der Fundamental-Matrix  $\mathbf{F}$  entspricht  $m = 7$ .

Die *Fitness* eines Chromosoms entscheidet über das Durchsetzungsvermögen innerhalb des Genpools. Die Berechnung erfordert eine direkte lineare Minimallösung für die entsprechende projektive Transformation (vgl. Abschnitt 4.1.5.2, 4.1.5.3 bzw. 4.3.4). Anschließend wird für alle Punkte der geometrische Fehler ermittelt und der robuste *Least-Median-of-Squares* (vgl. *LMedS* im Anhang A.3.2) als Maß für die Fitness verwendet. Damit entspricht ein kleiner Wert einem großen Durchsetzungsvermögen.

#### 4.4.1.2 Genetische Operatoren

Um eine existierende Population zu optimieren, können biologisch motivierte Veränderungsmechanismen eingesetzt werden. Das Vermischen der Erbanlagen von erfolgreichen Individuen wird mit der so genannten *Überkreuzung* (cross-over) realisiert. Dabei werden zwei Chromosomen an einer Stelle auseinander geschnitten und überkreuz wieder zusammengesetzt.

Die Durchführung der Operation und die Trennstelle werden dabei zufällig ausgewählt. Es ist allerdings darauf zu achten, dass beim Austausch von Teilabschnitten keine doppelten Gene entstehen. Ab der Trennstelle werden also nur diejenigen Genpaare berücksichtigt, die ein individuelles Vorkommen gewährleisten. Bei der Implementierung wurde die Wahrscheinlichkeit zur Durchführung einer Überkreuzung  $P_C$  mit 0,5 so gewählt, dass im Durchschnitt jedes zweite Chromosom verändert wird.



**Abb. 4.12:** Beispiele für genetische Operatoren

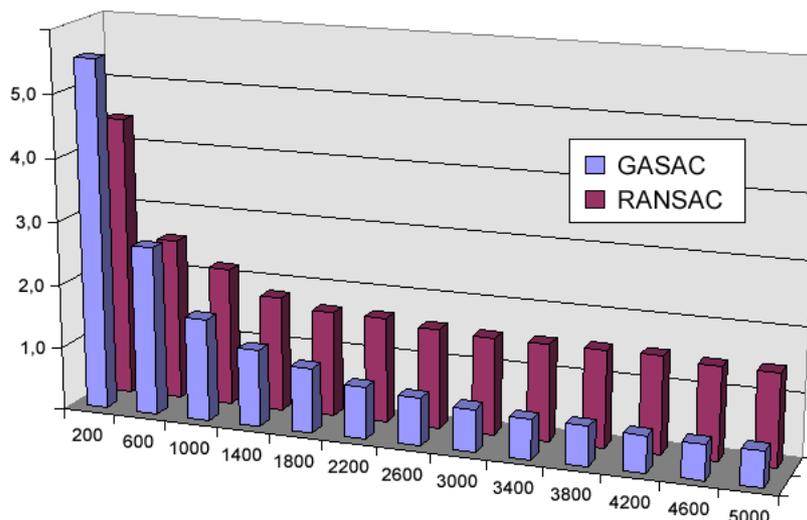
Um das Verharren in einem suboptimalen lokalen Minimum zu verhindern, muss neues Genmaterial zugeführt werden können. Dies wird durch zufälliges Verändern der Erbinformation erreicht, wobei auch hier darauf zu achten ist, dass keine doppelten Gene entstehen. Die Wahrscheinlichkeit für die *Mutation* eines Gens  $P_M$  wurde mit  $\frac{1}{2m}$  so gewählt, dass durchschnittlich bei jedem zweiten Chromosom ein Gen verändert wird.

Neben diesen beiden wichtigen Operatoren sind manchmal auch *Invertierungen* und *Rekombinationen* gebräuchlich [BRA91]. Da bei der gewählten Repräsentation aber die Reihenfolge der Gene keine Rolle spielt und die Länge der Chromosomen konstant bleiben soll, wurde auf diese zusätzlichen Operatoren verzichtet.

#### 4.4.1.3 Reproduktionsplan

Die Strategie für eine erfolgreiche evolutionäre Entwicklung ist im Reproduktionsplan verankert. Zunächst wird für die *Initialisierung* eine Population mit  $N$  Individuen erzeugt, wobei die Gene *zufällig* ausgewählt werden. Anschließend wird für jedes Chromosom die entsprechende Fitness bestimmt. Aus dieser Population müssen dann Elternpaare für die Reproduktion ausgewählt und mit Hilfe der genetischen Operatoren Nachkommen produziert werden. Dieser Vorgang wird solange fortgesetzt, bis  $M$  neue Individuen erzeugt wurden. Schließlich muss auch die Fitness der Nachkommen bestimmt und eine Auswahl der besten  $N$  Individuen erfolgen, um ein Anwachsen der gesamten Population zu verhindern. Dieser Reproduktionsschritt wird solange wiederholt, bis ein geeignetes Abbruchkriterium erfüllt ist.

Um die *Konvergenz* des genetischen Algorithmus zu verbessern, wurden einige Modifikationen vorgenommen. Bei einer geringen Populationsgröße können trotz einer ausgewogenen Initialisierung schnell Individuen mit identischen Genen entstehen. Das Entfernen dieser doppelten Individuen aus dem Genpool beschleunigt den Optimierungsprozess deutlich. Die aus den genetischen Operationen entstehenden Individuen sind nicht zwangsläufig immer besser als die ursprünglichen. Um sicherzustellen, dass sich eine erreichte Lösung nicht wieder verschlechtern kann, wird in jedem Zyklus das Chromosom mit der besten Fitness unverändert in den Genpool übernommen. Am Ende definiert das Individuum mit der maximalen Fitness und dem kleinsten LMedS-Fehler die Lösung vom genetischen Algorithmus.



**Abb. 4.13:** Verbesserung des geometrischen Fehlers [in Pixel] über 5.000 Versuche

#### 4.4.1.4 Vergleich der Methoden

Um den Unterschied zwischen zufälligem und systematischem Probieren zu demonstrieren, soll der GASAC-Ansatz dem RANSAC-Verfahren gegenübergestellt werden (siehe Anhang A.3.3). Beide Implementierungen verwenden für die Berechnung der Fundamental-Matrix die LMedS-Minimierung (Anhang A.3.2) des symmetrischen Epipolar-Abstands (Abschnitt 4.3.5). Für den Vergleich wurde eine Testumgebung mit 25 Bildpaaren der NOFRETETE in der größten Auflösung und jeweils 50 stark gestörten homologen Bildpunkten erzeugt.

Abb. 4.13 zeigt die Verbesserung des geometrischen Fehlers bei der robusten Parameterschätzung von  $\mathbf{F}$  in Abhängigkeit von der Anzahl der Versuche. Bei dem RANSAC-Verfahren wurden insgesamt 5.000 Iterationen und bei dem genetischen Algorithmus mit  $N = 200$  und  $M = 400$  in 12 Zyklen genauso viele Bewertungen der Fitness durchgeführt. Um ein statistisch repräsentatives Ergebnis zu erhalten, sind alle Bildpaare 100-mal ausgewertet und die Fehlerwerte arithmetisch gemittelt worden. Der Vorteil des GASAC-Verfahrens wird erst deutlich, wenn man die Anzahl der notwendigen Bewertungen bis zum Erreichen der optimalen Lösung betrachtet (vgl. Tab. 4.6).

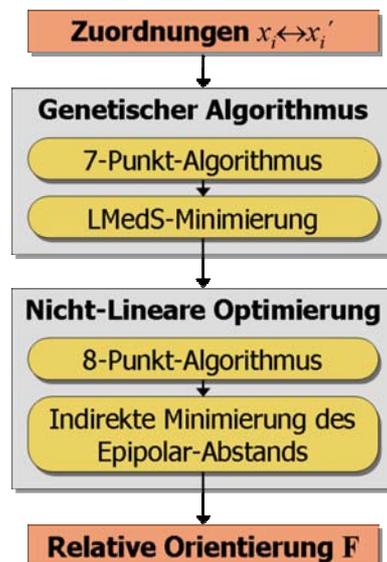
#### 4.4.2 Robustes Orientierungs-Verfahren

Mit den Erkenntnissen der vorangegangenen Abschnitte soll nun ein beschleunigtes Verfahren zur robusten Berechnung der Fundamental-Matrix  $\mathbf{F}$  aus mindestens sieben homologen Bildpunkten vorgeschlagen werden. Zunächst werden mit dem *genetischen Algorithmus* grobe Fehler durch falsche Zuordnungen beseitigt (Abschnitt 4.4.1). Für das systematische Probieren wird der schnelle *7-Punkt-Algorithmus* eingesetzt (Abschnitt 4.3.4) und die Lösung durch *LMedS-Minimierung* des symmetrischen Epipolar-Abstands gefunden.

Methode	$N$	$M$	Bewertungen
Theoretisch			ca. 100 Mio.
RANSAC			57.233
GASAC	50	100	7.052
	100	200	4.465
	<b>100</b>	<b>400</b>	<b>4.147</b>
	100	800	5.186
	200	400	5.633

**Tab. 4.6:** Anzahl der Bewertungen in Abhängigkeit von der Populationsgröße  $N$  und dem Fortpflanzungsverhalten

Die durch ungenaue Messungen entstandenen kleinen Fehler werden anschließend mit Hilfe der *nicht-linearen Optimierung* reduziert (Abschnitt 4.3.5). Falls nach dem Beseitigen der Ausreißer noch mehr als sieben Punkte optimiert werden müssen, empfiehlt sich der zusätzliche Einsatz des für überbestimmte Gleichungssysteme geeigneten *8-Punkt-Algorithmus* (Abschnitt 4.3.3). Die dadurch verbesserten Näherungswerte stabilisieren und beschleunigen den nachfolgenden Optimierungsprozess. Eine Zusammenfassung der vorgeschlagenen Methode zur robusten Bestimmung von  $\mathbf{F}$  ist in Abb. 4.14 dargestellt.



**Abb. 4.14:** Orientierung eines Bildpaares aus homologen Bildpunkten durch die robuste Berechnung der Fundamental-Matrix

Methode	$N$	Geom. Fehler [Pixel]		$\Delta F$		Epipol Fehler		CPU [in Sek.]
		$\emptyset$	Max	$\emptyset$	Max	$\emptyset$	Max	
<b>Normierter 8-Punkt-Algorithmus</b>								
Linear	350	2,152	19,770	0,219	1,249	0,499	1,255	0,01
Nicht-linear	350	2,131	19,185	0,219	1,249	0,449	0,966	2.516,10
<b>Robuste Schätzung mit GASAC (7-Punkt-Algorithmus)</b>								
M-Schätzer (TUKEY)	253	0,568	6,059	0,215	1,247	0,438	0,960	5,04
M-Schätzer (HUBER)	236	0,535	4,980	0,215	1,246	0,404	0,929	4,95
LMedS	231	0,493	4,820	0,211	1,245	0,369	0,788	4,53
Nicht-linear	232	0,474	4,599	0,210	1,245	0,368	0,788	684,24

Tab. 4.7: Vergleich der verschiedenen Methoden zur Bestimmung der Fundamental-Matrix

### 4.4.3 Experimentelle Ergebnisse

Um die unterschiedlichen Methoden zur Berechnung der Fundamental-Matrix  $\mathbf{F}$  vergleichen zu können, wurden 25 Bildpaare der NOFRETETE in der hohen Auflösung mit  $N = 350$  homologen Bildpunkten verwendet. Für ein statistisch repräsentatives Ergebnis sind jeweils 100 Testläufe ausgewertet worden (vgl. Tab. 4.7). Zusätzlich zum geometrischen Bildfehler (vgl. Abschnitt 4.3.5) sind die Abweichungen der berechneten Fundamental-Matrizen zu den durch klassische Kalibrierung bestimmten Matrizen  $\mathbf{F}_{Kalib}$  mit Hilfe der Frobenius-Norm  $\Delta F = \|\mathbf{F} - \mathbf{F}_{Kalib}\|$  dargestellt. Um schließlich auch eine visuelle Kontrolle der Ergebnisse zu ermöglichen, sind für zwei Beispiele einige Epipolar-Geraden überlagert worden (siehe Abb. 4.15 und Abb. 4.16).

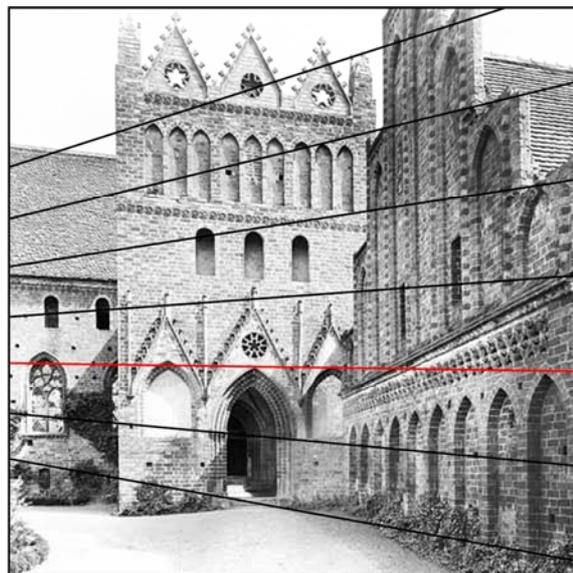
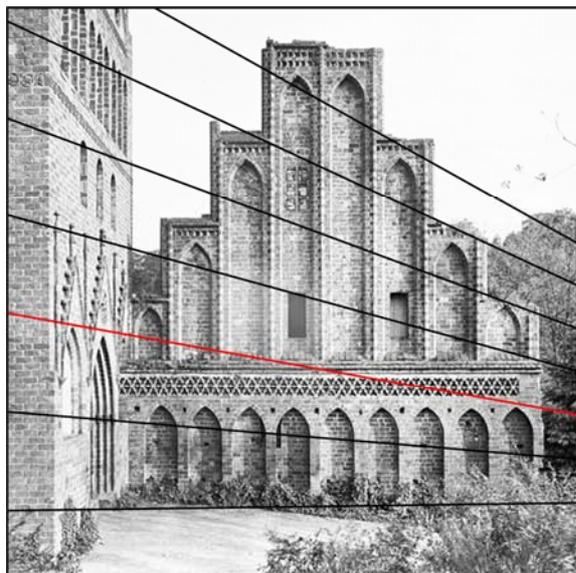


Abb. 4.15: Berechnete Epipolar-Geometrie für das Stereo-Bildpaar des Klosters in CHORIN



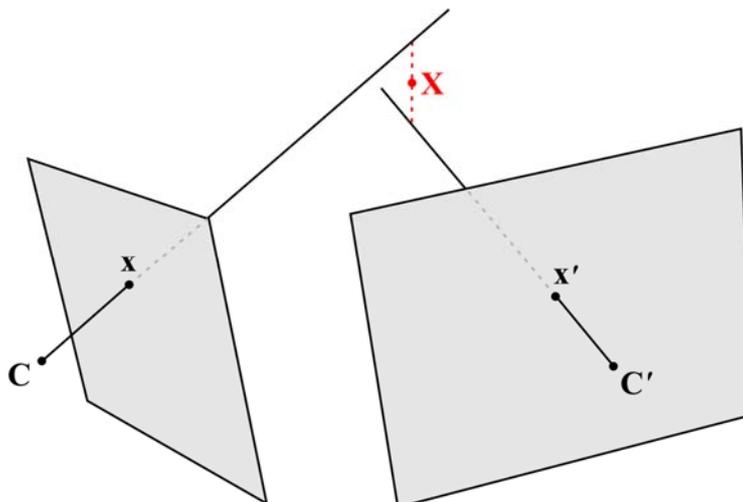
**Abb. 4.16:** Automatisch geschätzte Epipolar-Geometrie für die Bilder der Kirche in VALBONNE ohne die Modellierung der radialen Objektivverzeichnung

## 4.5 Triangulation

Nachdem die Epipolar-Geometrie mit Hilfe der Fundamental-Matrix bestimmt wurde, können anschließend die zugehörigen Projektions-Matrizen  $\mathbf{P}$  und  $\mathbf{P}'$  aus  $\mathbf{F}$  abgeleitet werden (vgl. Abschnitt 4.3.2). Die *lineare homogene Triangulation* in Abschnitt 4.5.1 berechnet durch einen räumlichen Vorwärtsschnitt den Objektpunkt  $\mathbf{X}$ , dessen Projektionen den gemessenen Bildpunkten  $\mathbf{x} \leftrightarrow \mathbf{x}'$  entsprechen

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad \text{und} \quad \mathbf{x}' = \mathbf{P}'\mathbf{X}. \quad (4.73)$$

Wenn die gemessenen Bildpunkte  $\mathbf{x}$  und  $\mathbf{x}'$  allerdings fehlerbehaftet sind, dann schneiden sich die Projektionsstrahlen nicht, d.h. es existiert kein Objektpunkt  $\mathbf{X}$ , dessen Projektionen die Epipolar-Bedingung  $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$  exakt erfüllen. In diesem Fall wird üblicherweise der Objektpunkt mit dem geringsten Abstand zu den beiden Projektionsstrahlen berechnet (vgl. Abb. 4.17).



**Abb. 4.17:** Triangulations-Problem

Diese Vorgehensweise ist in der projektiven Geometrie allerdings nicht zu empfehlen, da die Abstände im projektiven Raum  $\mathbb{P}^3$  von der gewählten Basis abhängig sind. In Abschnitt 4.5.2 soll daher eine verbesserte Triangulations-Methode beschrieben werden, die projektiv-invariant ist, da nur Abstände in den Bildern optimiert werden.

### 4.5.1 Lineare Triangulation

Mit Hilfe der *linearen homogenen Triangulation* können die räumlichen Koordinaten aus fehlerfreien homologen Bildpunkten  $\mathbf{x} \leftrightarrow \mathbf{x}'$  einfach berechnet werden (vgl. Abb. 4.20a). Da die beiden homogenen Projektions-Matrizen  $\mathbf{P}$  und  $\mathbf{P}'$  einen unterschiedlichen Skalierungsfaktor besitzen, muss dieser zunächst mit Hilfe des Kreuzproduktes

$$\mathbf{x} \times (\mathbf{P} \mathbf{X}) = \begin{bmatrix} x(\mathbf{p}^3 \mathbf{X}) - (\mathbf{p}^1 \mathbf{X}) \\ y(\mathbf{p}^3 \mathbf{X}) - (\mathbf{p}^2 \mathbf{X}) \\ x(\mathbf{p}^2 \mathbf{X}) - y(\mathbf{p}^1 \mathbf{X}) \end{bmatrix} = \mathbf{0} \quad (4.74)$$

entfernt werden, wobei  $\mathbf{p}^i$  die Zeile  $i$  von  $\mathbf{P}$  bezeichnet. Wenn jedes der beiden Bilder zwei linear unabhängige Gleichungen liefert, dann können die drei Unbekannten des Objektpunktes  $\mathbf{X}$  bestimmt werden. Durch Umformen erhält man das lineare homogene Gleichungssystem

$$\mathbf{A} \mathbf{X} = \mathbf{0} \quad \text{mit} \quad \mathbf{A} = \begin{bmatrix} x\mathbf{p}^3 - \mathbf{p}^1 \\ y\mathbf{p}^3 - \mathbf{p}^2 \\ x'\mathbf{p}'^3 - \mathbf{p}'^1 \\ y'\mathbf{p}'^3 - \mathbf{p}'^2 \end{bmatrix} \quad \text{und} \quad \mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}, \quad (4.75)$$

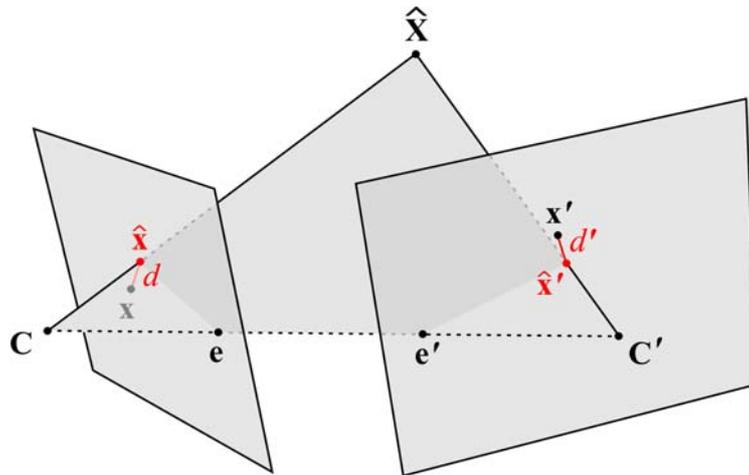
das mit Hilfe der Singulärwert-Zerlegung gelöst werden kann (vgl. Anhang A.2.1). Diese Triangulations-Methode lässt sich auch leicht verallgemeinern, wenn ein Punkt in mehr als zwei Bildern abgebildet ist (vgl. Abschnitt 5.1.3).

### 4.5.2 Optimale Triangulation

Sind die Bildpunkte  $\mathbf{x}$  und  $\mathbf{x}'$  durch ungenaue Messungen leicht gestört, dann sollte die verbesserte Triangulations-Methode von HARTLEY und STURM [HAR95] eingesetzt werden. Das Verfahren optimiert die Position der Bildpunkte bezüglich der Epipolar-Geometrie, indem das Minimum der Fehlerfunktion

$$\min_{\mathbf{x}, \mathbf{x}'} d(\mathbf{x}, \mathbf{l})^2 + d(\mathbf{x}', \mathbf{l}')^2 \quad (4.76)$$

gesucht wird, wobei  $d$  den senkrechten (euklidischen) Abstand des Punktes  $\mathbf{x}$  zu der entsprechenden Geraden  $\mathbf{l}$  im Bild bezeichnet (siehe Abb. 4.18). Durch die geschickte Formulierung des Minimierungsproblems reduziert sich die Aufgabe auf die Suche reeller Nullstellen eines Polynoms sechsten Grades.



**Abb. 4.18:** Optimale Triangulation durch Minimierung des geometrischen Fehlers im Bild

Unter der Annahme einer GAUSSschen Fehlerverteilung ist dieses Verfahren statistisch optimal. Um die verbesserten Positionen zu berechnen, muss zunächst jeder homologe Bildpunkt  $\mathbf{x} = (x, y, 1)^T$  und  $\mathbf{x}' = (x', y', 1)^T$  mit den Transformations-Matrizen

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & -x \\ 0 & 1 & -y \\ 0 & 0 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{T}' = \begin{bmatrix} 1 & 0 & -x' \\ 0 & 1 & -y' \\ 0 & 0 & 1 \end{bmatrix} \quad (4.77)$$

in den Ursprung verschoben und die Fundamental-Matrix  $\mathbf{F}$  an diese zentrierten Koordinaten angepasst werden

$$\mathbf{F}' = \mathbf{T}'^{-T} \mathbf{F} \mathbf{T}^{-1}. \quad (4.78)$$

Anschließend können die zugehörigen Epipole  $\mathbf{e} = (e_1, e_2, e_3)^T$  und  $\mathbf{e}' = (e'_1, e'_2, e'_3)^T$  aus  $\mathbf{F}'$  extrahiert werden (vgl. Abschnitt 4.3.2).

Jetzt müssen die Koordinatensysteme der beiden Bilder noch so gedreht werden, dass die Epipole jeweils auf der  $x$ -Achse liegen. Dazu werden die beiden Epipole mit einem Faktor normiert, so dass jeweils  $e_1^2 + e_2^2 = 1$  gilt. Schließlich können die Rotations-Matrizen

$$\mathbf{R} = \begin{bmatrix} e_1 & e_2 & 0 \\ -e_2 & e_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{R}' = \begin{bmatrix} e'_1 & e'_2 & 0 \\ -e'_2 & e'_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.79)$$

aufgestellt werden und nach der Transformation

$$\hat{\mathbf{F}} = \mathbf{R}' \mathbf{F}' \mathbf{R}^T \quad (4.80)$$

besitzt die resultierende Fundamental-Matrix  $\hat{\mathbf{F}}$  eine spezielle Form, deren geometrischer Zusammenhang in Abb. 4.19 dargestellt ist.

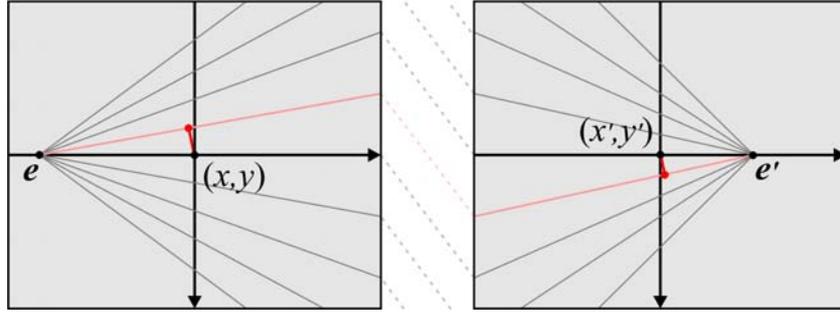


Abb. 4.19: Minimierungsaufgabe nach der Koordinatentransformation

Nach dieser Umformulierung besteht das Ziel der Optimierung darin, diejenige Epipolar-Ebene zu finden, deren Epipolar-Geraden den geringsten Abstand zum Ursprung haben. Dazu muss das Polynom

$$p(t) = t \left( \left( \hat{\mathbf{F}}_{22}t + \hat{\mathbf{F}}_{23} \right)^2 + e_3'^2 \left( \hat{\mathbf{F}}_{32}t + \hat{\mathbf{F}}_{33} \right)^2 \right)^2 - \left( \hat{\mathbf{F}}_{22}\hat{\mathbf{F}}_{33} - \hat{\mathbf{F}}_{23}\hat{\mathbf{F}}_{32} \right) \left( 1 + e_3^2 t^2 \right)^2 \left( \hat{\mathbf{F}}_{22}t + \hat{\mathbf{F}}_{23} \right) \left( \hat{\mathbf{F}}_{32}t + \hat{\mathbf{F}}_{33} \right) \quad (4.81)$$

vom Grad sechs gelöst werden. Durch Ausmultiplizieren nach der Variablen  $t$  ergibt sich die Form

$$p(t) = at^6 + bt^5 + ct^4 + dt^3 + et^2 + ft + g = 0,$$

wobei die Koeffizienten durch

$$\begin{aligned} a &= \hat{\mathbf{F}}_{22}\hat{\mathbf{F}}_{32}h e_3^4 \\ b &= \hat{\mathbf{F}}_{22}^4 + \hat{\mathbf{F}}_{23}^2\hat{\mathbf{F}}_{32}^2 e_3^4 + \hat{\mathbf{F}}_{32}^4 e_3'^4 + \hat{\mathbf{F}}_{22}^2 \left( -\hat{\mathbf{F}}_{33}^2 e_3^4 + 2\hat{\mathbf{F}}_{32}^2 e_3'^2 \right) \\ c &= 2\hat{\mathbf{F}}_{22}\hat{\mathbf{F}}_{32}h e_3^2 + \hat{\mathbf{F}}_{23}\hat{\mathbf{F}}_{33}h e_3^4 + 4 \left( \hat{\mathbf{F}}_{22}\hat{\mathbf{F}}_{23}\hat{\mathbf{F}}_{32}^2 e_3'^2 + \hat{\mathbf{F}}_{22}^2\hat{\mathbf{F}}_{32}\hat{\mathbf{F}}_{33} e_3'^2 + \hat{\mathbf{F}}_{32}^3\hat{\mathbf{F}}_{33} e_3'^4 + \hat{\mathbf{F}}_{22}^3\hat{\mathbf{F}}_{23} \right) \\ d &= 2 \left( 4\hat{\mathbf{F}}_{22}\hat{\mathbf{F}}_{23}\hat{\mathbf{F}}_{32}\hat{\mathbf{F}}_{33} e_3'^2 + \hat{\mathbf{F}}_{22}^2 \left( 3\hat{\mathbf{F}}_{23}^2 + \hat{\mathbf{F}}_{33}^2 \left( -e_3^2 + e_3'^2 \right) \right) + \hat{\mathbf{F}}_{32}^2 \left( 3\hat{\mathbf{F}}_{33}^2 e_3'^4 + \hat{\mathbf{F}}_{23}^2 \left( e_3^2 + e_3'^2 \right) \right) \right) \\ e &= \hat{\mathbf{F}}_{22}\hat{\mathbf{F}}_{23} \left( 4\hat{\mathbf{F}}_{23}^2 + \hat{\mathbf{F}}_{32}^2 - 2\hat{\mathbf{F}}_{33}^2 \left( e_3^2 - 2e_3'^2 \right) \right) + 2\hat{\mathbf{F}}_{32}\hat{\mathbf{F}}_{33} \left( 2\hat{\mathbf{F}}_{33}^2 e_3'^4 + \hat{\mathbf{F}}_{23}^2 \left( e_3^2 + 2e_3'^2 \right) \right) - \hat{\mathbf{F}}_{22}^2\hat{\mathbf{F}}_{32}\hat{\mathbf{F}}_{33} \\ f &= \hat{\mathbf{F}}_{23}^4 - \hat{\mathbf{F}}_{22}^2\hat{\mathbf{F}}_{33}^2 + \hat{\mathbf{F}}_{33}^4 e_3'^4 + \hat{\mathbf{F}}_{23}^2 \left( \hat{\mathbf{F}}_{32}^2 + 2\hat{\mathbf{F}}_{33}^2 e_3'^2 \right) \\ g &= \hat{\mathbf{F}}_{23}\hat{\mathbf{F}}_{33}h \end{aligned} \quad (4.82)$$

mit der Abkürzung  $h = \hat{\mathbf{F}}_{23}\hat{\mathbf{F}}_{32} - \hat{\mathbf{F}}_{22}\hat{\mathbf{F}}_{33}$  definiert sind.

Nach der numerischen Lösung des Polynoms z.B. mit ZRHQR [PRE02] erhält man sechs zum Teil komplexe Nullstellen. Es ist allerdings darauf zu achten, dass mehrere Koeffizienten sehr dicht bei 0 liegen können und sich der Grad des Polynoms dadurch reduzieren kann. Um diese Lösungsmenge wieder einzuschränken, muss die Funktion

$$s(t) = \frac{t^2}{1 + e_3^2 t^2} + \frac{\left( \hat{\mathbf{F}}_{32}t + \hat{\mathbf{F}}_{33} \right)^2}{\left( \hat{\mathbf{F}}_{22}t + \hat{\mathbf{F}}_{23} \right)^2 + e_3'^2 \left( \hat{\mathbf{F}}_{32}t + \hat{\mathbf{F}}_{33} \right)^2} \quad (4.83)$$

für die reellen Lösungen von  $p(t)$  sowie für den *asymptotischen Wert*

$$s(\infty) = \frac{1}{e_3^2} + \frac{\hat{\mathbf{F}}_{32}^2}{\hat{\mathbf{F}}_{22}^2 + e_3^2 \hat{\mathbf{F}}_{32}^2} \quad (4.84)$$

ausgewertet werden. Von diesen sieben möglichen Lösungen definiert das  $t$  mit den geringsten Kosten die optimale Epipolar-Ebene für den homologen Bildpunkt  $\mathbf{x} \leftrightarrow \mathbf{x}'$ .

Auf den entsprechenden Epipolar-Geraden

$$\mathbf{l} = (te_3, 1, -t)^T \quad \text{und} \quad \mathbf{l}' = \hat{\mathbf{F}}(0, t, 1)^T = \left( -e_3'(\hat{\mathbf{F}}_{32}t + \hat{\mathbf{F}}_{33}), \hat{\mathbf{F}}_{22}t + \hat{\mathbf{F}}_{23}, \hat{\mathbf{F}}_{32}t + \hat{\mathbf{F}}_{33} \right)^T \quad (4.85)$$

müssen noch die Positionen mit dem kürzesten senkrechten Abstand zum Ursprung bestimmt werden, wobei für eine allgemeine Gerade  $\mathbf{l} = (a, b, c)$  der dichteste Punkt durch

$$\tilde{\mathbf{x}} = (-ac, -bc, a^2 + b^2) \quad (4.86)$$

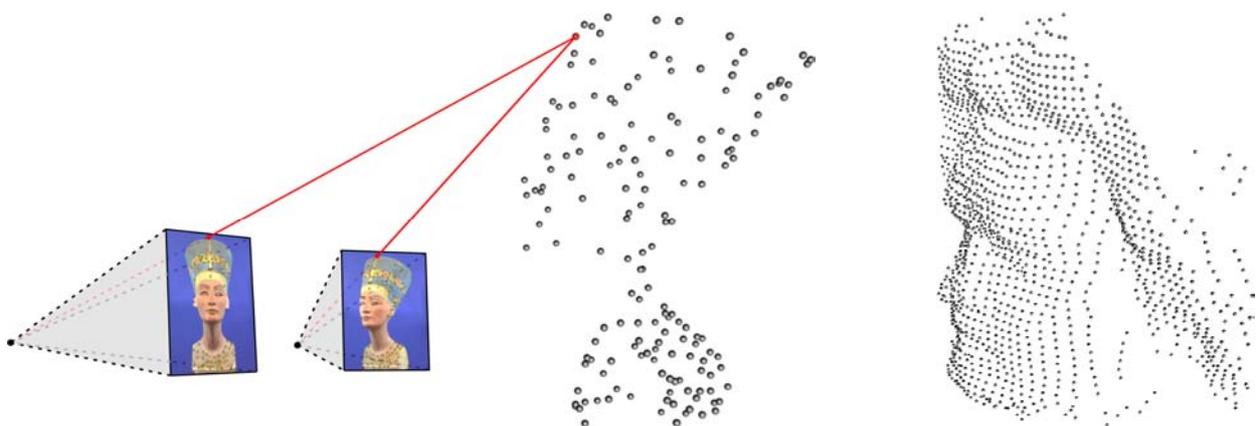
gegeben ist. Zum Abschluss werden die beiden optimierten Bildpunkte mit

$$\hat{\mathbf{x}} = \mathbf{T}^{-1} \mathbf{R}^T \tilde{\mathbf{x}} \quad \text{und} \quad \hat{\mathbf{x}}' = \mathbf{T}'^{-1} \mathbf{R}'^T \tilde{\mathbf{x}}' \quad (4.87)$$

wieder auf die ursprünglichen Koordinatensysteme zurück transformiert. Die verbesserten Bildpunkte  $\hat{\mathbf{x}}$  und  $\hat{\mathbf{x}}'$  erfüllen nun exakt die Epipolar-Bedingung  $\hat{\mathbf{x}}'^T \mathbf{F} \hat{\mathbf{x}} = 0$  und der gesuchte Objektpunkt  $\hat{\mathbf{X}}$  kann mit Hilfe der linearen homogenen Triangulation bestimmt werden (vgl. Abschnitt 4.5.1).

### 4.5.3 Experimentelle Ergebnisse

Um die Vorteile der optimalen Triangulation zu demonstrieren, sind synthetische Daten der Büste von BEETHOVEN ausgewertet worden (siehe Abb. 4.20b). Die Positionen von ca. 1.000 homologen Bildpunkten wurden mit einer GAUSSschen Normalverteilung gestört, wobei die Angabe 1 Prozent der Standardabweichung von 1 Pixel entspricht. Die resultierenden Abweichungen der triangulierten Objektpunkte von den bekannten Sollstellen sind in der Tab. 4.8 dargestellt.



a) Ergebnis der NOFRETETE mit den automatisch extrahierten homologen Interest-Punkten

b) Büste von BEETHOVEN mit idealen Zuordnungen

**Abb. 4.20:** Triangulation von Objektpunkten (in metrischer Darstellung)

Ansatz	Störung der Bildpunkte [in Prozent]						CPU [in Sek.]
	1		2		5		
	Ø	Max	Ø	Max	Ø	Max	
Lineare Triangulation	1,22	96,62	1,52	100,28	2,53	153,79	0,79
Optimale Triangulation	1,20	95,70	1,49	97,74	2,27	103,58	2,49

Tab. 4.8: Lagefehler der triangulierten Objektpunkte [in mm] bei verrauschten Bildpositionen

## 4.6 Diskussion

In diesem Kapitel konnte gezeigt werden, dass mit mindestens sieben homologen Bildpunkten eine projektive Rekonstruktion durch eine relative Orientierung des Bildpaares und eine anschließende Triangulation auch ohne Kenntnis der Kameraparameter möglich ist. Die direkte lineare Bestimmung der projektiven Transformationen mit Hilfe der DLT liefert bei einer geeigneten Normierung recht gute Ergebnisse. Es hat sich jedoch gezeigt, dass für die automatische Auswertung besonders robuste Schätzmethode unerlässlich sind. Die untersuchten Methoden zeigen die besten Resultate in Verbindung mit dem stabilen *LMedS*. Dieses Kriterium benötigt im Gegensatz zu den *M-Schätzern* auch keinen vordefinierten Schwellwert.

Bei der Bestimmung der relativen Orientierung kann der hohe Berechnungsaufwand ein wenig reduziert werden, indem der minimale *7-Punkt-Algorithmus* zur Bestimmung der Fundamental-Matrix verwendet wird. Eine signifikante Beschleunigung kann allerdings erreicht werden, wenn das zufällige Probieren durch eine systematische Strategie ersetzt wird. Dazu wurde in dieser Arbeit ein neues Verfahren mit Hilfe des genetischen Algorithmus *GASAC* vorgeschlagen. Während bei den meisten Optimierungsverfahren eine sequentielle Verbesserung erfolgt, wird bei den genetischen Algorithmen die Tatsache ausgenutzt, dass gleichzeitig mehrere bewertete Teillösungen parallel existieren. Diese Vorgehensweise ermöglicht es, die besten Parameterkombinationen zu extrahieren und zu besseren Lösungen zu kombinieren.

Des Weiteren wurde gezeigt, wie durch eine anschließende *nicht-lineare Optimierung* der algebraischen Lösungen kleine Messungenauigkeiten beseitigt werden können. Aus statistischer Sicht ist dieser Ansatz zwar optimal, aber der Berechnungsaufwand ist aufgrund der teilweise großen Parameteranzahl relativ hoch. Schließlich wurde für den räumlichen Vorwärtsschnitt ein Verfahren vorgestellt, das durch eine Optimierung der Bildpunkte bezüglich der Epipolar-Geometrie eine stabile und genaue Triangulation ermöglicht.

## 5 Projektive Mehrbild-Geometrie

Im vorangegangenen Kapitel wurden einzelne Stereo-Bildpaare projektiv orientiert und räumliche Koordinaten durch Triangulation der Bildpunkte bestimmt. Nun soll eine entsprechende Erweiterung für mehrere Bilder behandelt werden. Dabei wird zunächst erläutert, wie bei drei Bildern durch Ausnutzen der *Trifokal-Geometrie* die Anzahl der fehlerhaft zugeordneten Punktmerkmale weiter reduziert werden kann (Abschnitt 5.1). Danach werden im Abschnitt 5.2 zwei Ansätze für  $n \geq 3$  Bilder untersucht, die Näherungswerte für eine anschließende globale *Bündel-Ausgleichung* liefern (Abschnitt 5.3).

### 5.1 Orientierung eines Bildtripels

#### 5.1.1 Trifokal-Geometrie

Die *Trifokal-Geometrie* stellt eine Erweiterung der Epipolar-Geometrie für drei Bilder dar (vgl. Abschnitt 4.3.1). Wenn ein Objektpunkt in zwei Bildern identifiziert wurde, so kann seine geometrische Position im dritten Bild durch den Schnitt der entsprechenden Epipolar-Geraden vorhergesagt werden. Im Unterschied zum Bildpaar existiert beim Bildtripel jedoch ein eindeutiges Ergebnis, sofern der Punkt nicht in der Trifokal-Ebene der drei Projektionszentren liegt und die Projektionszentren nicht kollinear sind (siehe Abb. 5.1).

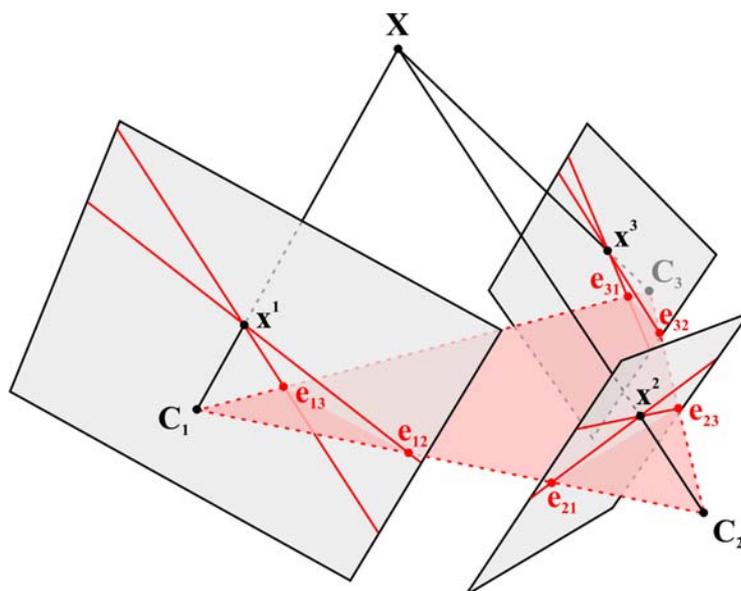


Abb. 5.1: Prinzip der Trifokal-Geometrie

Um die Trifokal-Geometrie ohne Kenntnis der inneren Orientierung aus homologen Bildpunkten zu bestimmen, wird zunächst der *minimale 6-Punkt-Algorithmus* vorgestellt (Abschnitt 5.1.2) und anschließend die *robuste Berechnung* mit GASAC erläutert (Abschnitt 5.1.3). Danach wird kurz auf die algebraische Repräsentation von projektiven Beziehungen zwischen Bildern mit Hilfe von *Mehrbild-Tensoren* eingegangen (Abschnitt 5.1.4).

### 5.1.2 Minimaler 6-Punkt-Algorithmus

CARLSSON und WEINSHALL [CAR98] formulierten ein Dualitätsprinzip, bei dem die Auffassung egal ist, ob Punkte von mehreren Kameras gesehen werden oder umgekehrt. Wenn es einen Algorithmus für  $m+4$  Punkte aus  $n$  Bildern gibt, dann existiert auch ein Algorithmus für  $n+4$  Punkte aus  $m$  Bildern. Demnach muss es zu dem 7-Punkt-Algorithmus für zwei Bilder (vgl. Abschnitt 4.3.4) auch einen 6-Punkt-Algorithmus für drei Bilder geben.

Das Verfahren von TORR und ZISSERMAN [TOR97] nach der Idee von QUAN [QUA94] berechnet aus sechs Punkten in allgemeiner Lage die Trifokal-Geometrie. Als Ergebnis erhält man bis zu drei Lösungen, genau wie bei dem minimalen 7-Punkt Algorithmus zur Bestimmung der Fundamental-Matrix. Zunächst muss in jedem der drei Bilder eine Normierung der Bildpunkte mit den Transformationen  $\mathbf{T}_i$  durchgeführt werden (vgl. Abschnitt 4.1.5.2). Dann können ohne Beschränkung der Allgemeinheit für die sechs Objektpunkte  $\mathbf{X}_j$  die projektiven Koordinaten  $(1,0,0,0)^T$ ,  $(0,1,0,0)^T$ ,  $(0,0,1,0)^T$ ,  $(0,0,0,1)^T$ ,  $(1,1,1,1)^T$  und  $(X,Y,Z,W)^T$  angenommen werden, wobei  $X$ ,  $Y$ ,  $Z$  und  $W$  unbekannt sind. Die zugehörigen sechs Punkte  $\mathbf{x}_j^i$  im Bild  $i=1,2,3$  können entsprechend durch die projektiven Koordinaten  $(1,0,0)^T$ ,  $(0,1,0)^T$ ,  $(0,0,1)^T$ ,  $(1,1,1)^T$ ,  $(x_5^i, y_5^i, w_5^i)^T$  und  $(x_6^i, y_6^i, w_6^i)^T$  definiert werden.

Um diese vereinfachte Darstellung zu erreichen, müssen die normierten Bildpunkte mit den  $3 \times 3$ -Homographie-Matrizen

$$\mathbf{A}_i = [\lambda_1 \mathbf{x}_1^i \quad \lambda_2 \mathbf{x}_2^i \quad \lambda_3 \mathbf{x}_3^i]^{-1}, \quad (\lambda_1, \lambda_2, \lambda_3)^T = [\mathbf{x}_1^i \quad \mathbf{x}_2^i \quad \mathbf{x}_3^i]^{-1} \mathbf{x}_4^i \quad (5.1)$$

transformiert werden. Anschließend gilt für jedes Bild der folgende Zusammenhang:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & x_5^i & x_6^i \\ 0 & 1 & 0 & 1 & y_5^i & y_6^i \\ 0 & 0 & 1 & 1 & w_5^i & w_6^i \end{bmatrix} = \begin{bmatrix} \alpha^i & 0 & 0 & \delta^i \\ 0 & \beta^i & 0 & \delta^i \\ 0 & 0 & \gamma^i & \delta^i \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & X \\ 0 & 1 & 0 & 0 & 1 & Y \\ 0 & 0 & 1 & 0 & 1 & Z \\ 0 & 0 & 0 & 1 & 1 & W \end{bmatrix} \quad (5.2)$$

Nun besteht die Aufgabe darin, die Koordinaten vom sechsten Objektpunkt  $(X,Y,Z,W)^T$  und die Parameter  $\alpha^i, \beta^i, \gamma^i, \delta^i$  der einzelnen Kameras zu rekonstruieren.

Durch Umformen der Gleichung (5.2) erhält man zunächst das lineare homogene Gleichungssystem

$$\begin{bmatrix} w_5^i & 0 & -x_5^i & w_5^i - x_5^i \\ 0 & w_5^i & -y_5^i & w_5^i - y_5^i \\ w_6^i X & 0 & -x_6^i Z & (w_6^i - x_6^i)W \\ 0 & w_6^i Y & -y_6^i Z & (w_6^i - y_6^i)W \end{bmatrix} \begin{pmatrix} \alpha^i \\ \beta^i \\ \gamma^i \\ \delta^i \end{pmatrix} = \mathbf{0} \quad \text{für } i = 1, 2, 3, \quad (5.3)$$

wobei die  $4 \times 4$ -Matrix auf der linken Seite singular mit dem Rang 3 ist [QUA94].

Verwendet man die Eigenschaft, dass die Determinante der Matrix 0 sein muss, so ergibt sich das folgende unterbestimmte Gleichungssystem in den Koordinaten der fünften und sechsten Bildpunkte:

$$\begin{bmatrix} x_5^1(w_6^1 - y_6^1) & y_5^1(x_6^1 - w_6^1) & w_5^1(y_6^1 - x_6^1) & w_6^1(y_5^1 - x_5^1) & y_6^1(x_5^1 - w_5^1) \\ x_5^2(w_6^2 - y_6^2) & y_5^2(x_6^2 - w_6^2) & w_5^2(y_6^2 - x_6^2) & w_6^2(y_5^2 - x_5^2) & y_6^2(x_5^2 - w_5^2) \\ x_5^3(w_6^3 - y_6^3) & y_5^3(x_6^3 - w_6^3) & w_5^3(y_6^3 - x_6^3) & w_6^3(y_5^3 - x_5^3) & y_6^3(x_5^3 - w_5^3) \end{bmatrix} \mathbf{t} = \mathbf{0} \quad (5.4)$$

Der Lösungsvektor  $\mathbf{t} = (WX - YZ, WY - YZ, WZ - YZ, XY - YZ, XZ - YZ)^T$  enthält den gesuchten Objektpunkt, der sich nach einer Singulärwert-Zerlegung (vgl. Anhang A.1) aus dem zweidimensionalen Nullraum berechnen lässt.

Seien  $\mathbf{u}$  bzw.  $\mathbf{v}$  die beiden rechten Nullvektoren, die jeweils dem Singulärwert 0 entsprechen, dann ist der gesuchte Vektor  $\mathbf{t}$  durch den Lösungsraum

$$\mathbf{t}_t = t\mathbf{u} + (1-t)\mathbf{v} \quad (5.5)$$

definiert, wobei  $t$  einen unbekanntes Skalar bezeichnet. Mit Hilfe der zusätzlichen Bedingung

$$t_1 t_2 t_5 - t_2 t_3 t_5 - t_2 t_4 t_5 = t_1 t_3 t_4 - t_2 t_3 t_4 - t_3 t_4 t_5 \quad (5.6)$$

in den Elementen von  $\mathbf{t} = (t_1, t_2, t_3, t_4, t_5)^T$  lässt sich diese unendliche Lösungsmenge wieder einschränken. Wie in Abschnitt 4.3.4 führt auch hier das Ausmultiplizieren zu einem kubischen Polynom in  $t$ . Nach der Berechnung der Nullstellen und dem Einsetzen in Gleichung (5.5) erhält man entweder eine oder drei reelle Lösungen für  $\mathbf{t}$ . Die gesuchten Koordinaten des sechsten Objektpunktes  $(X, Y, Z, W)^T$  lassen sich dann mit

$$\frac{X}{W} = \frac{t_4 - t_5}{t_2 - t_3}, \quad \frac{Y}{W} = \frac{t_4}{t_1 - t_3} \quad \text{und} \quad \frac{Z}{W} = \frac{t_5}{t_1 - t_2} \quad \text{für } W \neq 0 \quad (5.7)$$

aus  $\mathbf{t}$  extrahieren.

Im Anschluss können die Parameter  $\alpha^i, \beta^i, \gamma^i, \delta^i$  der einzelnen Kameras durch die Lösung der linearen homogenen Gleichungssysteme (5.3) bestimmt werden. Um die berechneten Projektions-Matrizen wieder an die Originaldaten anzupassen, müssen die initiale Transformation  $\mathbf{A}_i$  und die Normierung mit  $\mathbf{T}_i$  wieder aufgehoben werden

$$\mathbf{P}_i = \mathbf{T}_i^{-1} \mathbf{A}_i^{-1} \begin{bmatrix} \alpha^i & 0 & 0 & \delta^i \\ 0 & \beta^i & 0 & \delta^i \\ 0 & 0 & \gamma^i & \delta^i \end{bmatrix} \quad \text{für } i = 1, 2, 3. \quad (5.8)$$

Von den drei möglichen Lösungen wird schließlich diejenige mit dem geringsten geometrischen Fehler ausgewählt (siehe Gleichung 5.10). Wenn als Referenz die normierte Kamera  $\mathbf{P}_1 = [\mathbf{I} | \mathbf{0}]$  benötigt wird, dann können die Projektions-Matrizen mit

$$\mathbf{P}'_i = \mathbf{P}_i \mathbf{H}^{-1} \quad \text{und} \quad \mathbf{H} = \begin{bmatrix} \mathbf{P}_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{für } i = 1, \dots, m \quad (5.9)$$

entsprechend transformiert werden.

### 5.1.3 Robustes Orientierungs-Verfahren

Da in der Praxis häufig mehr als sechs Punkte verfügbar sind, kann durch ein systematisches Probieren mit dem genetischen Algorithmus GASAC eine Minimallösung gesucht werden, die durch möglichst viele Messwerte bestätigt wird (vgl. Abschnitt 4.4.1). Dazu muss für jede Auswahl der geometrische (euklidische) Abstand  $d$  der gemessenen homologen Punkte  $\mathbf{x}_j^1 \leftrightarrow \mathbf{x}_j^2 \leftrightarrow \mathbf{x}_j^3$  zu den idealen Bildpunkten berechnet werden, die durch Projektion der triangulierten Objektpunkte  $\mathbf{X}_j$  erzeugt werden

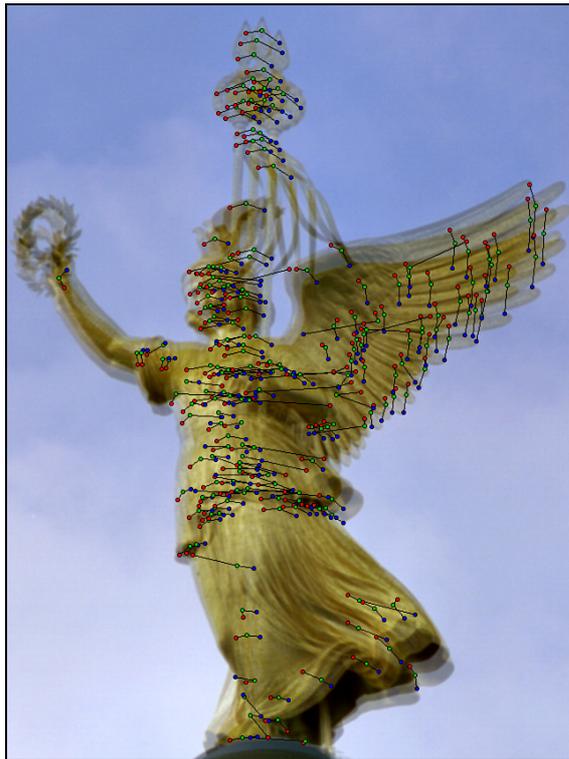
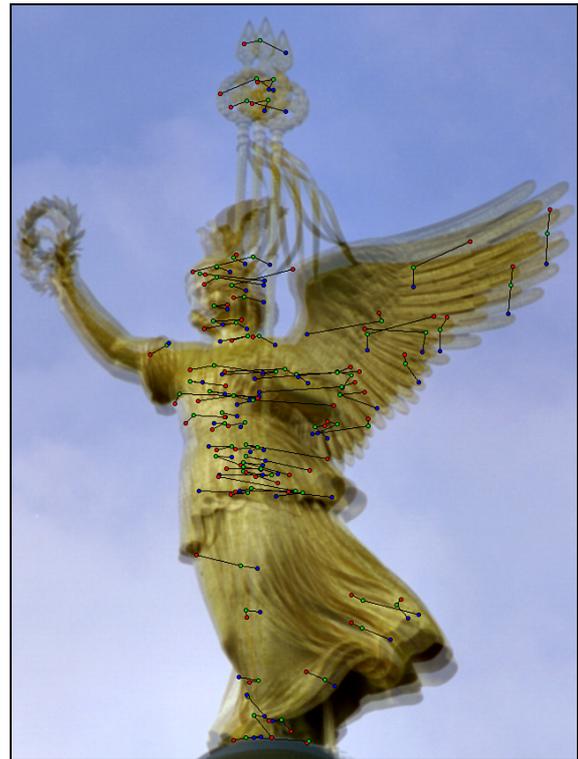
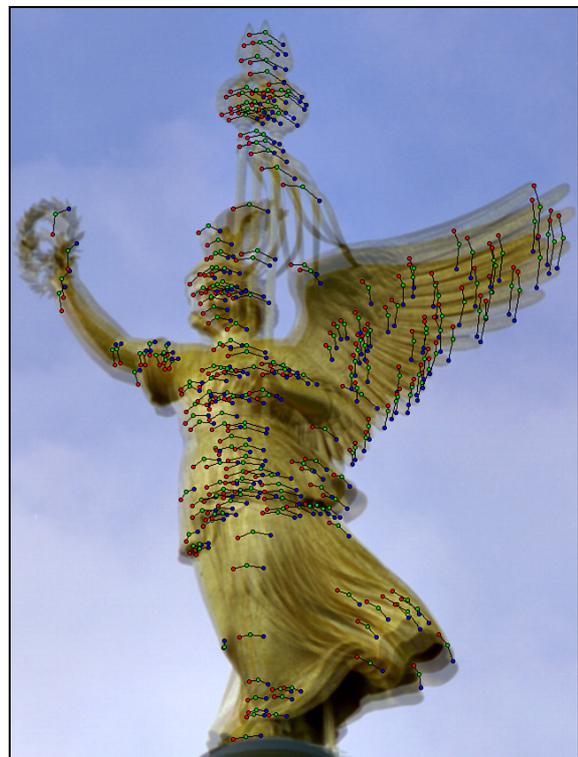
$$d(\mathbf{P}_1 \mathbf{X}_j, \mathbf{x}_j^1)^2 + d(\mathbf{P}_2 \mathbf{X}_j, \mathbf{x}_j^2)^2 + d(\mathbf{P}_3 \mathbf{X}_j, \mathbf{x}_j^3)^2. \quad (5.10)$$

Anschließend wird als Fehlerfunktion für die verschiedenen Teillösungen der LMedS minimiert (vgl. Anhang A.3.2).

Die Abb. 5.2 zeigt das Ergebnis einer robusten Berechnung der Trifokal-Geometrie am Beispiel der VIKTORIA-Statue. Von den ursprünglich 500 extrahierten Interest-Punkten wurden ca. 240 homologe Punkte in jedem Bildpaar identifiziert. Nach der Verkettung im Bildtripel verblieben noch 208 Punkte. Durch die Auswertung der Trifokal-Geometrie konnten 74 Fehlzuordnungen erkannt und aussortiert werden. Eine anschließende Korrespondenz-Analyse, bei der die Suchräume durch die Trifokal-Geometrie eingeschränkt wurden, steigerte die Anzahl der homologen Punkte auf 241.

### 5.1.4 Mehrbild-Tensoren

Der *Trifokal-Tensor*  $\mathcal{T}$  [SPE91] besitzt ähnliche Eigenschaften für drei Bilder wie die Fundamental-Matrix bei zweien. Er enthält die (projektiven) Beziehungen zwischen den Kameras unabhängig von dem abgebildeten Objekt. Obwohl der Trifokal-Tensor aus drei homogenen  $3 \times 3$ -Matrizen mit insgesamt 27 Elementen besteht, besitzt er nur 18 Freiheitsgrade.

**a)** Aus Bildpaaren verkettete Zuordnungen**b)** Erkannte Ausreißer**c)** Konsistente homologe Punkte**d)** Vervollständigtes Ergebnis

**Abb. 5.2:** Beispiel zur Reduktion von Zuordnungsfehlern durch die Auswertung der Trifokal-Geometrie. Dabei wurden drei Bilder der VIKTORIA-Statue überlagert und die homologen Punkte mit einer Linie verbunden.

Bilder	Tensor	Elemente	DOF	Punktzahl	
				Linear	Nicht-linear
2	$\mathbf{F}$	9	7	8	7*
3	$\mathcal{T}$	27	18	7	6*
4	$\mathcal{Q}$	81	29	6	6

**Tab. 5.1:** Übersicht der Mehrbild-Tensoren und die notwendige Punktzahl für eine Lösung (\* Mehrfach-Lösungen möglich)

Jede Projektions-Matrix  $\mathbf{P}_i$  enthält 12 Komponenten und  $\mathcal{T}$  ist invariant unter einer projektiven Transformation, die 16 Komponenten besitzt. Da aufgrund der homogenen Koordinaten die globale Skalierung irrelevant ist, ergeben sich für die projektiven Beziehungen zwischen drei Bildern insgesamt  $3 \times 11 - 15 = 18$  Freiheitsgrade. Während bei der Fundamental-Matrix eine zusätzliche Bedingung das Schneiden der Epipolar-Geraden in einem Punkt erzwingt (vgl. Abschnitt 4.3.3), so sind die fehlenden acht Bedingungen zwischen den Elementen des Trifokal-Tensors noch Gegenstand aktueller Forschungen [FÖR00].

Mit den drei Projektions-Matrizen  $\mathbf{P}_1 = [\mathbf{I} | \mathbf{0}]$ ,  $\mathbf{P}_2 = [a_{ij}]$  und  $\mathbf{P}_3 = [b_{ij}]$  aus dem 6-Punkt-Algorithmus (vgl. Abschnitt 5.1.2) kann der Trifokal-Tensor folgendermaßen bestimmt werden:

$$\mathcal{T}_{ijk} = a_{ji}b_{k4} - a_{j4}b_{ki} \quad \text{für } i, j, k = 1, 2, 3 \quad (5.11)$$

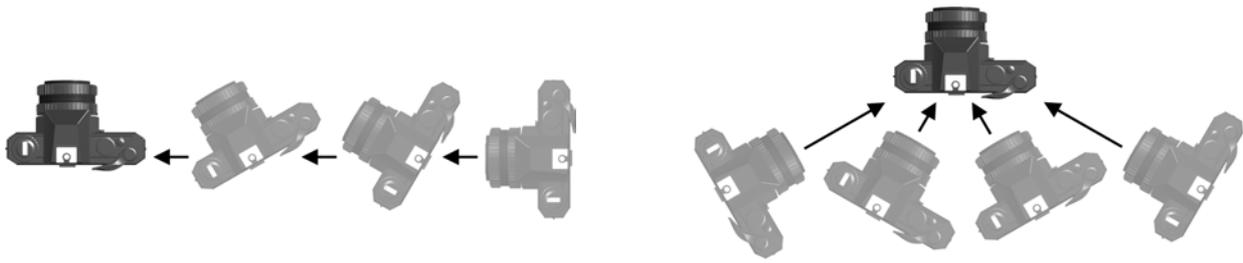
Der *Quadrifokal-Tensor*  $\mathcal{Q}$  [TRI95] stellt die konsequente Erweiterung für vier Bilder dar, doch ist er in der Praxis weniger gebräuchlich. Einige Eigenschaften der Mehrbild-Tensoren sind in Tab. 5.1 zusammengefasst.

## 5.2 Anpassen der projektiven Basis

Bislang wurden die homogenen Projektions-Matrizen für ein Stereo-Bildpaar nur bis auf einen unbekanntem Skalierungsfaktor bestimmt. Wenn man diesen Zusammenhang für mehrere Bildpaare explizit formuliert, dann gilt bei  $j$  betrachteten Punkten in  $i$  Bildern

$$\lambda_{ij} \mathbf{x}_j^i = \mathbf{P}_i \mathbf{X}_j. \quad (5.12)$$

Dieser Faktor  $\lambda$ , der die projektive Basis der Rekonstruktion definiert, wird häufig auch als *projektive Tiefe* bezeichnet [STU96]. Eine Verknüpfung von Bildpaaren kann aber erst dann erfolgen, wenn alle Rekonstruktionen eine gemeinsame projektive Tiefe  $\lambda$  besitzen. Die Aufgabe besteht also darin, die einzelnen projektiven Rekonstruktionen auf eine Referenz-Basis zu transformieren.



a) Sequentielle Anpassung der projektiven Rekonstruktionen auf eine Referenz-Basis

b) Parallele Anpassung der Rekonstruktionen auf einen gemeinsamen Schwerpunkt

**Abb. 5.3:** Kombination der projektiven Rekonstruktionen mit Hilfe einer räumlichen Homographie und einer projektiven Faktorisation

Während die *räumliche Homographie* in Abschnitt 5.2.1 eine sequentielle Verknüpfung durchführt, ist das Ziel der *projektiven Faktorisation* in Abschnitt 5.2.2 eine parallele Anpassung auf einen gemeinsamen Schwerpunkt (vgl. Abb. 5.3).

### 5.2.1 Räumliche Homographie

Die Überlappung von unterschiedlichen Rekonstruktionen kann mit Hilfe einer *Homographie im projektiven Raum*  $\mathbb{P}^3$  erreicht werden. Hierfür eignet sich das in Abschnitt 4.1.5.3 beschriebene Verfahren. Zunächst wird ein Bildpaar als Referenz für die projektive Basis definiert. Anschließend kann für alle anderen projektiven Rekonstruktionen eine Homographie  $\mathbf{H}$  bestimmt werden, die alle Objektpunkte auf die entsprechenden Referenzpunkte transformiert.

Diese Homographie muss ebenfalls auf die zugehörigen Projektions-Matrizen angewendet werden

$$\mathbf{X}'_j = \mathbf{H} \mathbf{X}_j \quad \text{und} \quad \mathbf{P}'_i = \mathbf{P}_i \mathbf{H}^{-1}. \quad (5.13)$$

Für eine Überlappung sind jeweils fünf gemeinsame Objektpunkte  $\mathbf{X}_j \leftrightarrow \mathbf{X}'_j$  notwendig. Durch die Bestimmung der Fundamental-Matrix stehen allerdings schon mindestens sieben Objektpunkte zur Auswahl (vgl. Abschnitt 4.3.4).

### 5.2.2 Projektive Faktorisation

Von STURM und TRIGGS [STU96] wurde eine parallele Anpassung der projektiven Rekonstruktionen vorgeschlagen. Als Voraussetzung müssen die Objektpunkte allerdings in allen Bildern erkennbar sein. Um die gesuchte Transformation zu bestimmen, müssen zunächst alle im Bild  $i = 1, \dots, m$  sichtbaren Punkte  $\mathbf{x}$  mit einer metrischen Transformation  $\mathbf{T}_i$  normiert werden, so dass der Schwerpunkt der Koordinaten im Ursprung liegt und der durchschnittliche geometrische Abstand vom Ursprung  $\sqrt{2}$  beträgt (vgl. Abschnitt 4.1.5.2). Anschließend kann durch Ausformulierung von Gleichung (5.12) die  $3m \times n$ -Gewichtungs-Matrix

$$\mathbf{W} = \begin{bmatrix} \lambda_{11}\hat{\mathbf{x}}_1 & \lambda_{12}\hat{\mathbf{x}}_2 & \cdots & \lambda_{1n}\hat{\mathbf{x}}_n \\ \lambda_{21}\hat{\mathbf{x}}_1 & \lambda_{22}\hat{\mathbf{x}}_2 & \cdots & \lambda_{2n}\hat{\mathbf{x}}_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m1}\hat{\mathbf{x}}_1 & \lambda_{m2}\hat{\mathbf{x}}_2 & \cdots & \lambda_{mn}\hat{\mathbf{x}}_n \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \vdots \\ \mathbf{P}_m \end{bmatrix} [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n] \quad (5.14)$$

aufgestellt werden, wobei  $\hat{\mathbf{x}}_j^i$  den normierten  $j$ -ten Punkt im Bild  $i$  und  $\lambda_{ij}$  die entsprechende projektive Tiefe bezeichnen. Liegen alle Objektpunkte  $\mathbf{X}_j$  in einem annähernd konstanten Abstand, dann können die gesuchten Faktoren zunächst mit einem Startwert  $\lambda_{ij}^{(0)} = 1$  grob geschätzt werden. Anschließend werden die unbekanntenen Skalierungsfaktoren durch ein relativ einfaches iteratives Verfahren bestimmt. Das Prinzip basiert auf einer geeigneten *Normierung der Zeilen und Spalten* der Gewichtungsmatrix  $\mathbf{W}$ .

Zuerst erfolgt eine Anpassung aller Spalten auf den Betrag 1, indem jede Spalte  $j$  durch den Faktor

$$\alpha_j = \sqrt{\sum_{i=1}^{3m} W_{ij}^2} \quad \text{für } j = 1, \dots, n \quad (5.15)$$

dividiert wird. Anschließend können die Zeilen von  $\mathbf{W}$  auf den Betrag 1 normiert werden, indem die zu einem homogenen Bildpunkt  $\hat{\mathbf{x}} = (x, y, w)^T$  gehörenden drei Zeilen durch den gemeinsamen Faktor

$$\beta_i = \sqrt{\sum_{j=1}^n \sum_{k=3i-2}^{3i} W_{kj}^2} \quad \text{für } i = 1, \dots, m \quad (5.16)$$

geteilt werden. Diese Veränderung der Zeilen hat allerdings wieder einen negativen Einfluss auf die zuvor angepassten Spaltenwerte.

Damit die einzelnen Beträge möglichst dicht bei 1 liegen, ist eine Wiederholung der Spalten- und Zeilen-Normierung empfehlenswert

$$\lambda_{ij}^{(l)} = \frac{\lambda_{ij}^{(l-1)}}{\alpha_j \beta_i} \quad \text{für } l = 1, \dots, 3, \quad (5.17)$$

wobei aufgrund der schnellen Konvergenz drei Iterationen in der Regel schon ausreichend sind.

Nachdem die Gewichtungsmatrix  $\mathbf{W}$  iterativ normiert wurde, muss sie mit der Singulärwert-Zerlegung faktorisiert werden

$$\mathbf{W} = \mathbf{U}\mathbf{D}\mathbf{V}^T. \quad (5.18)$$

Anschließend wird eine *Rang-4 Approximation* durchgeführt, indem nur die vier größten Singulärwerte berücksichtigt und die restlichen Diagonalelemente gelöscht werden.

Dadurch ergibt sich die reduzierte Diagonal-Matrix

$$\tilde{\mathbf{D}} = \text{diag}(d_1, d_2, d_3, d_4) \quad \text{mit } d_1 \geq d_2 \geq d_3 \geq d_4. \quad (5.19)$$

Der Einfluss auf die beteiligten Matrizen ist in Abb. 5.4 dargestellt.

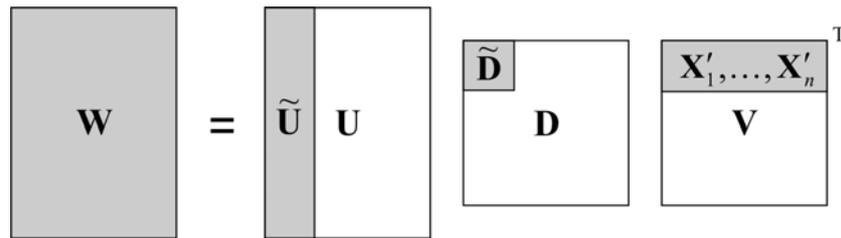


Abb. 5.4: Rang-4 Approximation bei der projektiven Faktorisierung

Die bezüglich eines gemeinsamen Skalierungsfaktors  $\lambda$  transformierten Projektions-Matrizen  $\hat{\mathbf{P}}_i$  und die neuen projektiven Objektpunkte  $\mathbf{X}'_j$  lassen sich durch

$$\left[ \hat{\mathbf{P}}_1, \hat{\mathbf{P}}_2, \dots, \hat{\mathbf{P}}_m \right]^T = \tilde{\mathbf{U}} \tilde{\mathbf{D}} \quad \text{und} \quad \left[ \mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_n \right] = \tilde{\mathbf{V}}^T \quad (5.20)$$

extrahieren. Zum Schluss muss die Normierung der Bildpunkte noch rückgängig gemacht werden, indem die einzelnen Projektions-Matrizen wieder an die ursprünglichen Koordinaten angepasst werden. Dies geschieht durch die Multiplikation mit den inversen Transformations-Matrizen

$$\mathbf{P}'_i = \mathbf{T}_i^{-1} \hat{\mathbf{P}}_i \quad \text{für} \quad i = 1, \dots, m. \quad (5.21)$$

Durch den Schwerpunkt wird allerdings eine willkürliche projektive Transformation eingeführt und das Referenzsystem verändert.

Um wieder die normierte Kamera  $\mathbf{P}'_1 = [\mathbf{I} | \mathbf{0}]$  zu erhalten, kann mit der Gleichung (5.9) eine räumliche Homographie  $\mathbf{H}$  berechnet werden, die auf alle Projektions-Matrizen  $\mathbf{P}'_i$  und Objektpunkte  $\mathbf{X}'_j$  angewendet werden muss (siehe Gleichung 5.13).

### 5.3 Projektive Bündel-Ausgleichung

Nachdem die einzelnen projektiven Rekonstruktionen angepasst wurden, liegen in der Regel gute Näherungswerte für eine nicht-lineare *projektive Bündel-Ausgleichung* vor. Im Prinzip spannt jedes Bild zwischen dem zugehörigen Projektionszentrum und den gemessenen Bildpunkten ein räumliches Strahlenbündel auf.

Verknüpft man die aus mehreren Bildern entstandenen Strahlenbündel, so entsteht ein dichtes Netz von Raumstrahlen, das bei geeigneter Aufnahmeconfiguration eine hohe geometrische Stabilität besitzt.

Die numerische Anpassung der Strahlenbündel zwischen einem Projektionszentrum und der Menge aller Objektpunkte, bzw. zwischen einem Objektpunkt und der Menge aller Projektionszentren, wird *Bündel-Ausgleichung* genannt (vgl. Abb. 5.5). Weiterführende Informationen sind z.B. in TRIGGS u.a. [TRI99], BARTOLI und STURM [BAR01] oder auch SCHUMACHER [SCH03] zu finden.

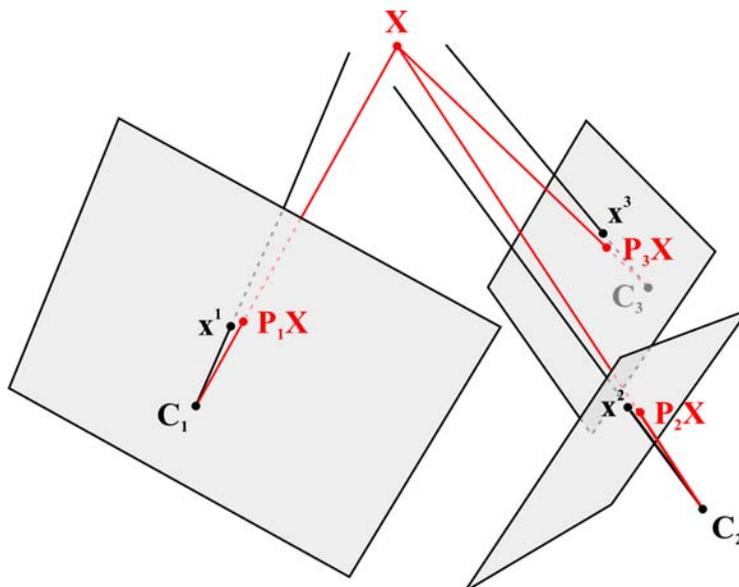


Abb. 5.5: Prinzip der Bündel-Ausgleichung

Für eine globale Optimierung wird das Minimum der Summe über alle  $m$  Punkte und  $n$  Bilder gesucht

$$\min_{\mathbf{P}_i, \mathbf{X}_j} \sum_{i=1}^n \sum_{j=1}^m d(\mathbf{P}_i \mathbf{X}_j, \mathbf{x}_j^i)^2, \quad (5.22)$$

wobei  $d(\mathbf{x}, \mathbf{y})$  den geometrischen (euklidischen) Abstand zwischen den homogenen Bildpunkten  $\mathbf{x}$  und  $\mathbf{y}$  bezeichnet. Diese Definition stellt eine Verallgemeinerung des Verfahrens aus Abschnitt 4.3.5 dar, welches die Fundamental-Matrix für ein Bildpaar durch indirekte Parametrisierung optimiert.

Da auch hier die normierte Kamera  $\mathbf{P}_1$  als Referenz konstant bleiben soll, erfolgt die nicht-lineare Optimierung über die restlichen  $(n-1)$  Projektions-Matrizen  $\mathbf{P}_i$  mit ihren jeweils 12 Elementen. Des Weiteren gehen noch  $m$  Objektpunkte  $\mathbf{X}_j$  mit jeweils drei Freiheitsgraden in die Berechnung ein, so dass für die Bündel-Ausgleichung insgesamt  $(n-1) \cdot 12 + m \cdot 3$  Parameter optimiert werden. Die Erweiterung der projektiven Rekonstruktion aus Kapitel 4 für die Mehrbild-Geometrie ist in Abb. 5.6 zusammengefasst.

## 5.4 Experimentelle Ergebnisse

Um die Verfahren zur Anpassung der projektiven Basis vergleichen zu können, wurden jeweils vier benachbarte Bilder der NOFRETETE-Bildsequenz ausgewertet. Bei minimaler Überlappung ergeben sich für die insgesamt 25 Bilder sieben Versuche, deren Ergebnisse arithmetisch gemittelt wurden. Die paarweisen Rekonstruktionen basieren auf jeweils 50 homologen Bildpunkten. Der geometrische Bildfehler bei der Kombination der drei Stereo-Bildpaare zu einem gemeinsamen Verband ist in Tab. 5.2 dargestellt.



Abb. 5.6: Übersicht der projektiven Rekonstruktion

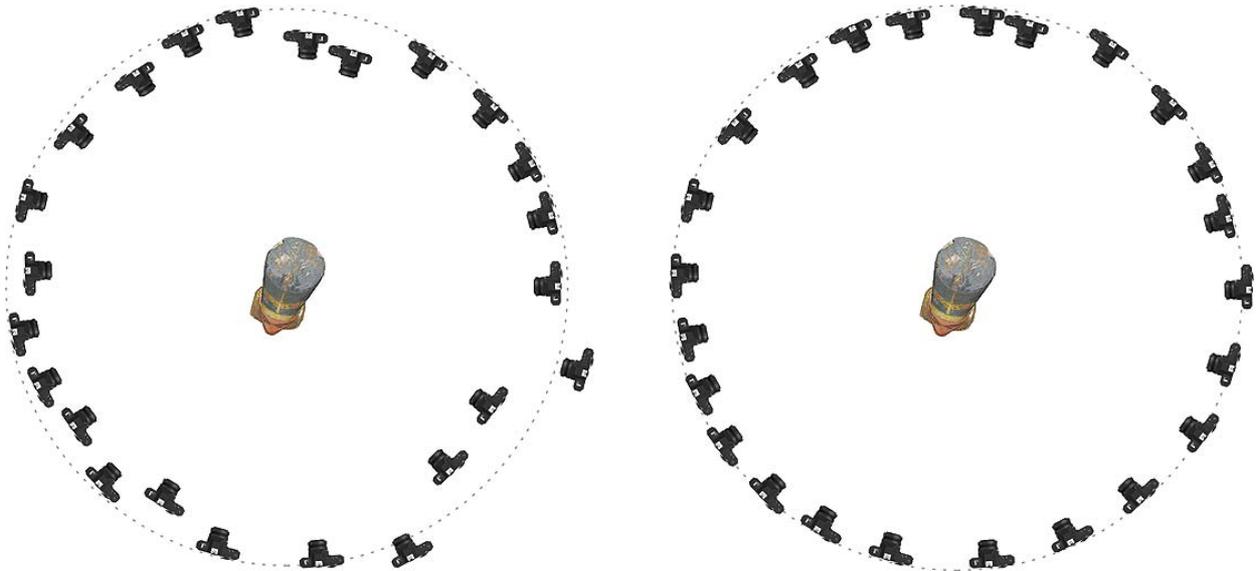
Die Auswirkung der Optimierung des Bildfehlers auf die Genauigkeit der äußeren Orientierung ist in Abb. 5.7 dargestellt. Das Anwendungsbeispiel zeigt die automatisch rekonstruierte Aufnahmeconfiguration der NOFRETETE-Sequenz von oben betrachtet. Um eine anschauliche Darstellung zu erhalten, ist die projektive Rekonstruktion mit Hilfe der Auto-Kalibrierung aus Kapitel 6 in den metrischen Raum transformiert worden.

## 5.5 Diskussion

Die robuste Bestimmung der Trifokal-Geometrie mit Hilfe des minimalen 6-Punkt-Algorithmus in Verbindung mit dem genetischen Algorithmus GASAC ist sehr zu empfehlen, da hiermit die Anzahl der fehlerhaft zugeordneten Punktmerkmale deutlich reduziert werden kann.

Ansatz	Mittlerer Fehler [in Pixeln]				Maximaler Fehler [in Pixeln]				CPU [in Sek.]
	1 ↔ 2	1 ↔ 3	1 ↔ 4	∅	1 ↔ 2	1 ↔ 3	1 ↔ 4	Max	
Ohne Anpassung	1,03	62,30	151,8	71,71	2,83	99,29	267,8	267,8	-
Projektive Faktorisierung	2,92	2,90	3,89	3,24	8,02	8,05	10,52	8,86	0,1
Räumliche Homographie	1,03	1,57	2,49	1,70	2,83	2,97	3,99	3,99	4,3
Bündel-Ausgleichung	0,67	0,72	0,66	0,68	1,34	1,49	1,36	1,49	56,9

Tab. 5.2: Vergleich der Methoden zur Anpassung der projektiven Basis bei drei Bildpaaren. Dargestellt ist der symmetrische Epipolar-Abstand und der Berechnungsaufwand.



a) Näherungswerte nach dem Anpassen der projektiven Basis

b) Ergebnis der Rekonstruktion nach der globalen Bündel-Ausgleichung

**Abb. 5.7:** Verbesserung der Rekonstruktion durch eine Bündel-Ausgleichung am Beispiel der NOFRETETE (in metrischer Darstellung). Durch die manuelle Drehung der Büste ist der Kameraabstand zwar konstant, die Betrachtungswinkel jedoch ein wenig unregelmäßig.

Die Verknüpfung der einzelnen projektiven Rekonstruktionen mit Hilfe einer *räumlichen Homographie* liefert bereits mit dem linearen Ansatz gute Ergebnisse. Die nicht-lineare Optimierung der Homographie zeigt kaum signifikante Verbesserungen. Obwohl Abstände im projektiven Raum minimiert werden, eignet sich dieses Verfahren zur Beschaffung von Näherungswerten für die Bündel-Ausgleichung. Durch den sequentiellen Ansatz wächst der Fehler allerdings mit der Anzahl der verwendeten Bilder.

Die *projektive Faktorisierung* verhält sich ein wenig robuster, da durch den parallelen Ansatz der Fehler gleichmäßiger verteilt wird. Nachteilig ist allerdings, dass die Konvergenz des Iterations-Verfahrens nicht garantiert und die Schätzung des Startwertes nur für relativ konstante Entfernungen erlaubt ist, was eine große Einschränkung für den Nahbereich bedeutet.

Die *Bündel-Ausgleichung* sollte generell als abschließender Schritt einer Rekonstruktion durchgeführt werden. Im projektiven Fall ist die Minimierung des geometrischen Fehlers im Bild statistisch optimal. Sie erfordert allerdings eine gute Initialisierung und bei einer großen Anzahl von Bildern oder Punkten wird der Optimierungsprozess sehr aufwändig.

## 6 Automatische Selbst-Kalibrierung

Durch Fortschritte im Verständnis der projektiven Geometrie ist es möglich geworden, verschiedene Informationen aus einer unkalibrierten Sequenz von Bildern zu extrahieren. Bereits im Jahr 1913 entwickelte KRUPPA [KRU13] Polynomgleichungen, aus denen die innere Orientierung abgeleitet werden konnte. Aber erst 1992 wurde die Idee wieder aufgegriffen und MAYBANK [MAY92] zeigte, dass bei konstanter innerer Orientierung eine automatische Selbst-Kalibrierung mit mindestens drei Bildern möglich ist.

Das Bedürfnis einer photogrammetrischen 3D-Rekonstruktion in Fällen, wo der klassische Kalibrierungsprozess zu aufwändig oder gar unmöglich ist, führte zu einer Reihe von Ansätzen zur automatischen Selbst-Kalibrierung. Diese Verfahren können prinzipiell in zwei Grundklassen eingeteilt werden:

- **Klasse A**

Die unbekannte Kalibrierungs-Matrix  $\mathbf{K}$  wird mit Hilfe der *Fundamental-Matrizen* aus der rekonstruierten Epipolar-Geometrie einzelner Bildpaare gewonnen [ZEL96, HAR97b, LUO97, MEN99]. Diese Methode setzt allerdings für alle Kameras eine konstante innere Orientierung voraus (vgl. Abschnitt 6.2). Des Weiteren wird nicht die Eigenschaft ausgenutzt, dass in der gesamten Bildsequenz die ideale Ebene  $\Pi_\infty$  identisch ist.

- **Klasse B**

Besser funktionieren Algorithmen, die  $\mathbf{K}$  aus der *projektiven Rekonstruktion* der Szene bestimmen [HEY97, TRI97, POL99b, POL99c]. Sie bieten die Möglichkeit, die innere Orientierung mit variierenden Parametern zu rekonstruieren (siehe Abschnitt 6.3). Ein weiterer Vorteil dieser Ansätze liegt in der Tatsache, dass weniger degenerierte Situationen existieren als bei den Verfahren der Klasse A [STU97, TRI97]. Da die einzelnen projektiven Rekonstruktionen vorher auf eine gemeinsame Basis transformiert werden müssen, ist allerdings der Berechnungsaufwand etwas höher.

Ein umfassende Zusammenfassung der verschiedenen Rekonstruktionsansätze ist z.B. bei FAUGERAS und LUONG [FAU01], FUSIELLO [FUS00], HARTLEY und ZISSERMAN [HAR00] sowie POLLEFEYS [POL99b] zu finden.

Zunächst sollen in Abschnitt 6.1 die Verwendung zusätzlicher Informationen erläutert und zweckmäßige Einschränkungen des Kamera-Modells diskutiert werden.

Nach einer kurzen Einführung in das Prinzip der Auto-Kalibrierung (Abschnitt 6.2) werden anschließend die Grundgedanken von zwei klassischen Methoden skizziert (Abschnitt 6.3). Zum Schluss wird das Verfahren zur Auto-Kalibrierung mit dem dualen absoluten Quadric vorgestellt und durch Eigenschaften der essentiellen Matrix verbessert (Abschnitt 6.4).

## 6.1 Verwendung zusätzlicher Informationen

Um von einer projektiven Rekonstruktion zu einer metrischen Darstellung zu gelangen, werden neben den Bildern noch zusätzliche Informationen benötigt (siehe Abb. 6.1). Dabei existieren im Prinzip die folgenden Möglichkeiten:

- **Direkte Rekonstruktion:** Das grundlegende Theorem der projektiven Geometrie besagt, dass die Kenntnis von fünf *Passpunkten* in allgemeiner Lage ausreicht, um die unbekannte  $4 \times 4$ -Matrix  $\mathbf{H}$  zu bestimmen, welche die Transformation vom projektiven zum euklidischen Raum durchführt [SEM52]. Durch die Berechnung der räumlichen Homographie  $\mathbf{H}$  können die euklidischen Objektpunkte und Projektions-Matrizen direkt mit

$$\mathbf{X}_E = \mathbf{H}\mathbf{X} \quad \text{bzw.} \quad \mathbf{P}_E = \mathbf{P}\mathbf{H}^{-1} \quad (6.1)$$

erzeugt werden (vgl. Abschnitt 4.1.5.3).

- **Modell-basierte Rekonstruktion:** Eine weitere Möglichkeit besteht in der Ausnutzung von bekannten Objekteigenschaften, wie z.B. *Parallelitäten*, *Winkeln* und *Streckenlängen* [CRI99]. Diese Ansätze sind aufgrund des benötigten Spezialwissens über die abgebildeten Objekte allerdings nicht universell einsetzbar.

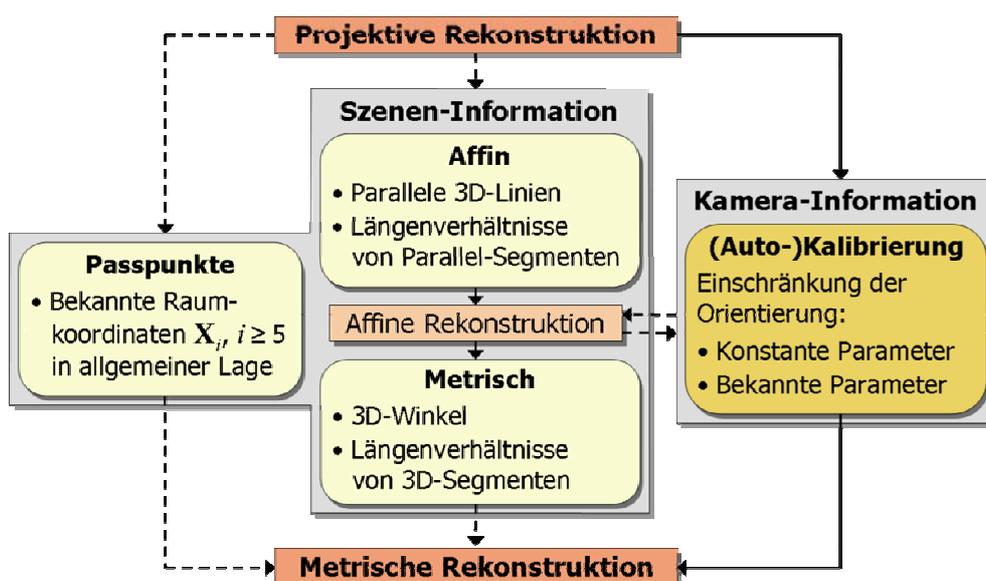


Abb. 6.1: Verwendung von Zusatzinformation für eine metrische Rekonstruktion

- **Einschränkung der äußeren Orientierung:** Für *besondere Aufnahmekonfigurationen* (z.B. Drehteller) oder *spezielle Kamerabewegungen* lassen sich zusätzliche Bedingungen formulieren, die ebenfalls eine metrische Rekonstruktion erlauben [FIT98].
- **Einschränkung der inneren Orientierung:** Schließlich können die fehlenden Informationen auch durch Annahmen über bekannte oder konstante *Parameter der inneren Orientierung* gewonnen werden. Da diese flexible Methode die potentiellen Anwendungsmöglichkeiten am wenigsten einschränkt, bildet sie die Grundlage für das Verfahren zur Auto-Kalibrierung.

Von einem allgemeinen Standpunkt betrachtet, besitzt eine projektive Rekonstruktion 15 Freiheitsgrade. Die angestrebte metrische Rekonstruktion hat mit drei Rotationen, drei Translationen und einer Skalierung insgesamt sieben Freiheitsgrade, so dass noch  $15 - 7 = 8$  Bedingungen benötigt werden.

Wenn bei  $n$  Bildern  $b$  Parameter der inneren Orientierung *bekannt* sind, dann ergeben sich  $n \cdot b$  Bedingungen. Falls  $k$  Parameter als *konstant* angenommen werden können, ergibt das  $(n-1) \cdot k$  Bedingungen. Eine metrische Rekonstruktion ist also möglich, wenn die Ungleichung

$$n \cdot b + (n-1) \cdot k \geq 8 \quad (6.2)$$

erfüllt ist. Für diese unterschiedlichen Modell-Annahmen sind einige typische Beispiele in Tab. 6.1 zusammengefasst. Dabei wird deutlich, dass bei einem *bekanntem Bildhauptpunkt* eine direkte lineare Lösung mit mindestens drei Bildern möglich ist. Diese Eigenschaft wird bei dem implementierten Verfahren zur Auto-Kalibrierung in Abschnitt 6.4 genutzt.

Annahme	Orientierungs-Parameter					Bildanzahl
	$\alpha_x$	$\alpha_y$	$x_0, y_0$	$\gamma$	$s$	
Scherung und Seitenverhältnis konstant	-	-	-	K	K	5*
Alle Parameter von <b>K</b> konstant	K	K	K	K	K	3
Scherung bekannt	-	-	-	-	B	8*
Scherung und Seitenverhältnis bekannt	-	-	-	B	B	4*
Bildhauptpunkt bekannt	-	-	B	-	-	4 (5 linear)
Scherung und Bildhauptpunkt bekannt	-	-	B	-	B	3 linear
Nur Kamerakonstante unbekannt	-	-	B	B	B	2 (3 linear)
Alle Parameter von <b>K</b> bekannt	B	B	B	B	B	2

**Tab. 6.1:** Anzahl der benötigten Bilder bei unterschiedlichen Modell-Annahmen [HAR00]  
(B = Bekannt, K = Konstant, \* Mehrfach-Lösungen möglich)

## 6.2 Prinzip der Auto-Kalibrierung

Um die Grundidee der Auto-Kalibrierung zu veranschaulichen, sollen die gesuchten acht Parameter geometrisch gedeutet werden. Drei dieser Parameter beschreiben die äußere Orientierung und werden durch die *ideale Ebene* definiert. Mit deren Kenntnis kann eine affine Rekonstruktion erreicht werden (Abschnitt 6.2.1).

Die Projektion des *absoluten Kegelschnitts* bestimmt die verbleibenden fünf Parameter der inneren Orientierung und ermöglicht dadurch eine metrische Rekonstruktion (Abschnitt 6.2.2).

### 6.2.1 Affine Rekonstruktion mit der idealen Ebene

Bei einer projektiven Abbildung schneiden sich im Allgemeinen parallele Geraden in so genannten *Fluchtpunkten*. Die ideale Ebene  $\Pi_\infty$  (vgl. Abschnitt 4.1.4.3) wird durch drei dieser Fluchtpunkte definiert (siehe Abb. 6.2). Eine Transformation, welche die ideale Ebene in die unendlich ferne Position  $\Pi_\infty = [0, 0, 0, 1]^T$  verschiebt, erzeugt aus einer projektiven eine *affine Rekonstruktion* (vgl. Abschnitt 4.1.1), in der parallele Linien und Ebenen wieder hergestellt sind.

HARTLEY [HAR98] beschreibt ein Verfahren, mit dem die Lage der idealen Ebene  $\Pi_\infty$  grob geschätzt werden kann. Das Prinzip besteht in dem Ausnutzen der *orientierten projektiven Geometrie* [STO91]. Dabei werden zunächst die Objektpunkte durch das Vorzeichen so ausgerichtet, dass sie alle vor der Projektionsebene  $\mathcal{R}$  liegen. Anschließend werden alle Projektions-Matrizen durch das Vorzeichen auf die orientierten Punkte gedreht.

Durch die Maximierung eines Systems von Ungleichungen (*Cheiral Inequalities*) mit Hilfe einer für negative Variablen modifizierten *linearen Programmierung* [PRE02] kann schließlich eine *quasi-affine Rekonstruktion* erreicht werden.

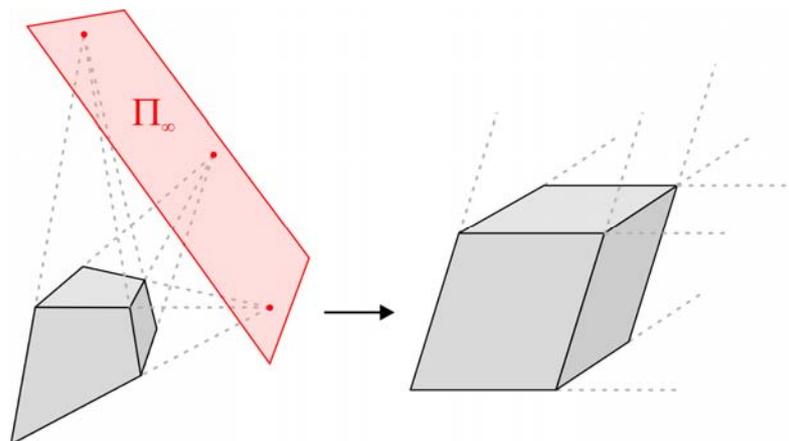


Abb. 6.2: Affine Rekonstruktion mit der idealen Ebene [POL99b]

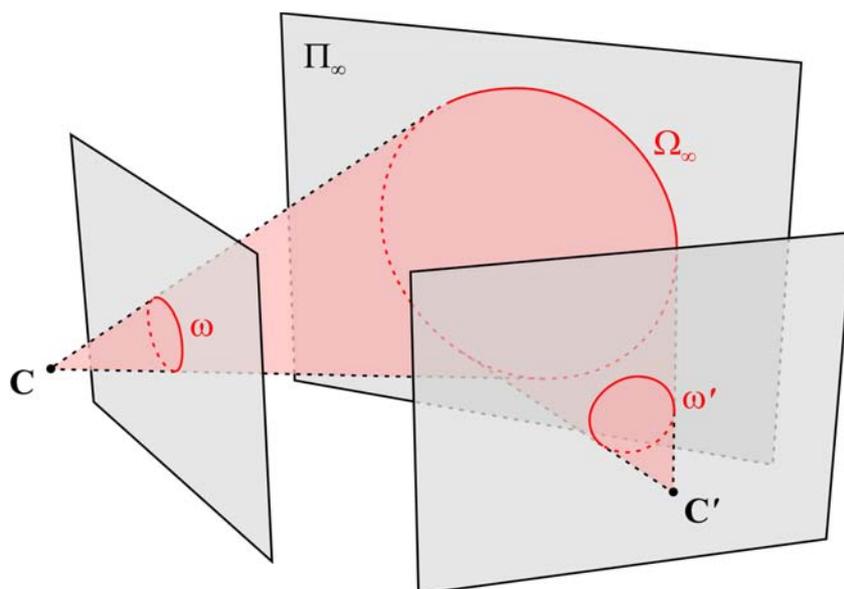
LUONG [LUO96] sowie POLLEFEYS u.a. [POL99c] schlagen für die Berechnung der idealen Ebene  $\Pi_\infty$  ein recht aufwändiges Polynomsystem vor, das auch als *Modulus Constraint* bekannt ist. Allerdings wird als Voraussetzung für alle Kameras eine konstante innere Orientierung angenommen. In dem Abschnitt 6.4 wird erläutert, wie die ideale Ebene  $\Pi_\infty$  durch die Berechnung des dualen absoluten Quadric wesentlich eleganter gefunden werden kann.

### 6.2.2 Innere Orientierung durch den absoluten Kegelschnitt

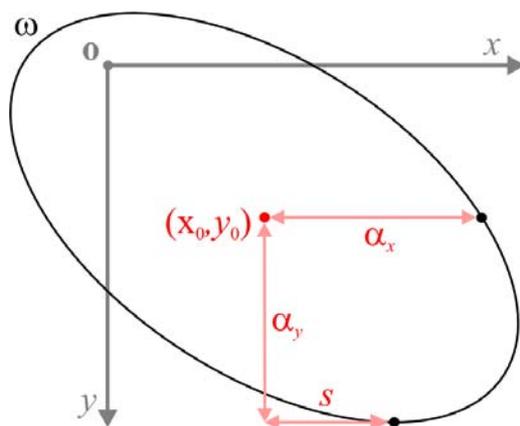
Eines der bedeutendsten Konzepte der Selbst-Kalibrierung ist der *absolute Kegelschnitt*  $\Omega_\infty$  (vgl. Abschnitt 4.1.4.4) und seine *Projektion*  $\omega$  in den Bildern (Image of the Absolute Conic). Da die Form des Kegelschnitts bezüglich einer metrischen Transformation invariant ist, bleibt die Projektion bei einer Bewegung der Kamera unverändert und ist somit unabhängig von der äußeren Orientierung. Die fünf Freiheitsgrade des abgebildeten Kegelschnitts  $\omega$  stehen also nur in Beziehung zu den fünf Parametern der inneren Orientierung.

Um diesen Zusammenhang zu veranschaulichen, wird zunächst eine *normierte Kamera*  $\mathbf{P}_N = [\mathbf{I} | -\mathbf{C}]$  betrachtet (vgl. Tab. 4.4). Die Projektion des absoluten Kegelschnitts  $\omega$  stellt in diesem Fall einen Kreis mit dem Radius  $r = \sqrt{-1}$  dar. Bei einer *projektiven Kamera* mit der Kalibrierungs-Matrix

$$\mathbf{K}_p = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.3)$$



**Abb. 6.3:** Der absolute Kegelschnitt  $\Omega_\infty$  in der idealen Ebene  $\Pi_\infty$  mit seinen Projektionen  $\omega$  und  $\omega'$



**Abb. 6.4:** Einfluss der Orientierungs-Parameter auf die Projektion des absoluten Kegelschnitts  $\omega$

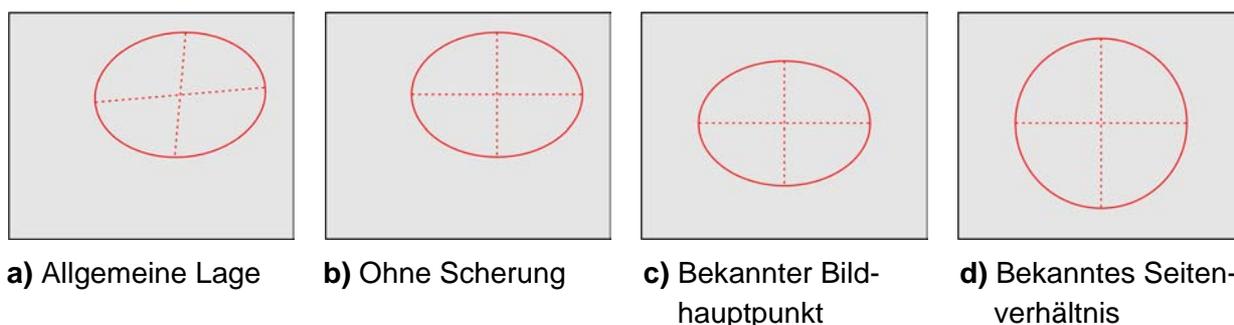
ist die Projektion  $\omega$  ein Kegelschnitt mit den *imaginären* Punkten

$$\frac{(x-x_0)^2}{\alpha_x^2} + \frac{(y-y_0)^2}{\alpha_y^2} + 2 \frac{(x-x_0)(y-y_0)\cos\theta}{\alpha_x\alpha_y} + 1 = 0 \quad \text{für} \quad \theta = -\operatorname{arccot}\left(\frac{s}{\alpha_x}\right). \quad (6.4)$$

Die geometrischen Beziehungen zwischen den Parametern der inneren Orientierung in  $\mathbf{K}$  und der Projektion  $\omega$  des absoluten Kegelschnitts sind in Abb. 6.4 skizziert. Dabei ist der Scherungs-Parameter  $s$  durch die relative  $x$ -Koordinate des höchsten Punktes von  $\omega$  definiert. Als Beispiel sind in Abb. 6.5 die imaginären Punkte der projizierten Kegelschnitte für verschiedene Modell-Annahmen visualisiert (vgl. Tab. 6.1).

### 6.3 Klassische Methoden zur Auto-Kalibrierung

Im Folgenden sollen zwei Grundgedanken der klassischen Auto-Kalibrierung aufgegriffen werden. Die häufig verwendeten *Kruppa-Gleichungen* (Abschnitt 6.3.1) und die Eigenschaft *identischer Singulärwerte* der *essentiellen Matrix* (Abschnitt 6.3.2) zählen zu den Verfahren der Klasse A und setzen eine konstante innere Orientierung voraus. In der dargestellten Form ist für die Bestimmung der Kalibrierungs-Matrix  $\mathbf{K}$  eine nicht-lineare Optimierung notwendig, für die gute Näherungswerte erforderlich sind.



**a)** Allgemeine Lage

**b)** Ohne Scherung

**c)** Bekannter Bildhauptpunkt

**d)** Bekanntes Seitenverhältnis

**Abb. 6.5:** Verschiedene Projektionen  $\omega$  des absoluten Kegelschnitts mit zunehmenden Einschränkungen der Orientierungs-Parameter [POL99b]

### 6.3.1 Die KRUPPA-Gleichungen

Durch die folgende Verknüpfung der Kalibrierungs-Matrix

$$\boldsymbol{\omega}^* = \mathbf{K}\mathbf{K}^T \quad (6.5)$$

erhält man für jedes Bild eine symmetrische und positiv definierte  $3 \times 3$  Matrix mit fünf Freiheitsgraden, die häufig auch als *KRUPPA-Matrix* bezeichnet wird. Sie entspricht der Projektion des absoluten Kegelschnitts in der dualen Repräsentation (vgl. Abschnitt 6.2.2). Mit Hilfe der so genannten *KRUPPA-Gleichungen*

$$[\mathbf{e}']_x \boldsymbol{\omega}^* [\mathbf{e}']_x = \mathbf{F} \boldsymbol{\omega}^* \mathbf{F}^T \quad (6.6)$$

kann ein Zusammenhang zwischen der berechneten Fundamental-Matrix  $\mathbf{F}$  und den gesuchten Parametern von  $\boldsymbol{\omega}^*$  definiert werden [MAY92, FAU92]. Für eine konstante innere Orientierung gilt  $\boldsymbol{\omega}^* = \boldsymbol{\omega}^{*'}$  und die schiefsymmetrische Matrix mit dem Epipol  $\mathbf{e}'$  lässt sich direkt aus  $\mathbf{F}$  bestimmen (vgl. Abschnitt 4.3.2). Jede Fundamental-Matrix liefert zwei unabhängige quadratische Gleichungen in den Koeffizienten von  $\boldsymbol{\omega}^*$ .

Geometrisch entsprechen diese Bedingungen zwei Epipolar-Ebenen, die tangential zum absoluten Kegelschnitt  $\Omega_\infty$  liegen (siehe Abb. 6.6). Wenn  $\mathbf{F}$  für jede Bildkombination berechnet wird, ist eine nicht-lineare Lösung bereits ab drei Bildern möglich, die nach FAUGERAS [FAU96] allerdings recht instabil ist.

HARTLEY [HAR97b] schlägt daher die Berechnung der KRUPPA-Gleichungen mit der Singulärwert-Zerlegung vor (vgl. Anhang A.1). Sei

$$\mathbf{F} = \mathbf{U} \text{diag}(\lambda, \mu, 0) \mathbf{V}^T \quad \text{mit} \quad \mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3] \quad \text{und} \quad \mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3] \quad (6.7)$$

die SVD einer Fundamental-Matrix  $\mathbf{F}$ , dann sind die KRUPPA-Gleichungen durch

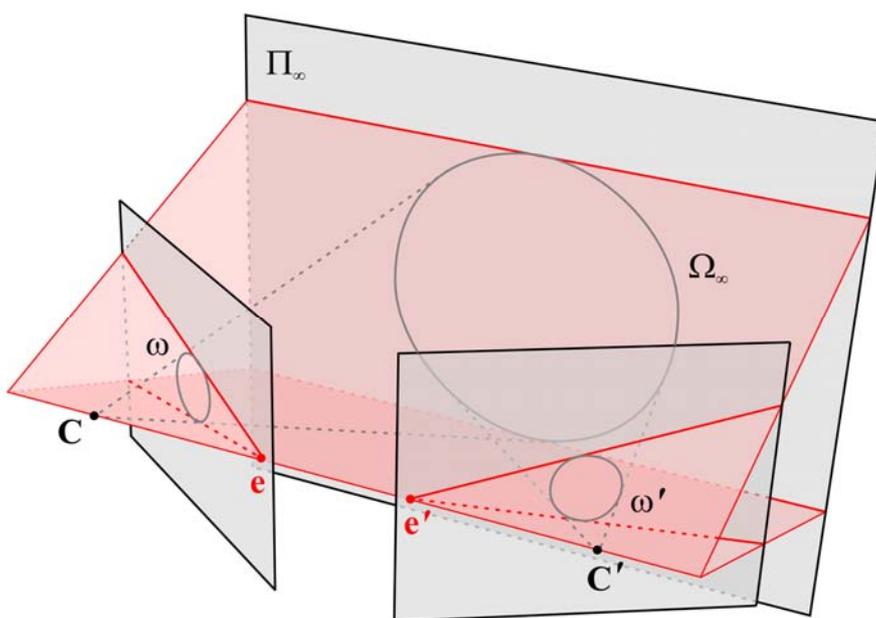


Abb. 6.6: Geometrische Deutung der KRUPPA-Gleichungen

$$\frac{\mathbf{u}_1^T \boldsymbol{\omega}^* \mathbf{u}_1}{\mu^2 \mathbf{v}_2^T \boldsymbol{\omega}^* \mathbf{v}_2} = \frac{\mathbf{u}_2^T \boldsymbol{\omega}^* \mathbf{u}_2}{\lambda^2 \mathbf{v}_1^T \boldsymbol{\omega}^* \mathbf{v}_1} = -\frac{\mathbf{u}_1^T \boldsymbol{\omega}^* \mathbf{u}_2}{\lambda \mu \mathbf{v}_1^T \boldsymbol{\omega}^* \mathbf{v}_2} \quad (6.8)$$

gegeben. Das Ziel der Optimierung besteht nun darin, die drei Ausdrücke in Gleichung (6.8) durch Anpassen der Elemente von  $\boldsymbol{\omega}^*$  möglichst gut anzugleichen. Da nur zwei der drei Ausdrücke unabhängig sind, kann bei  $n$  Bildern die folgende Fehlerfunktion minimiert werden:

$$\min_{\boldsymbol{\omega}^*} \sum_{i=1}^{\binom{n}{2}} w_i \left( \left( \frac{\mathbf{u}_{1i}^T \boldsymbol{\omega}^* \mathbf{u}_{1i}}{\mu_i^2 \mathbf{v}_{2i}^T \boldsymbol{\omega}^* \mathbf{v}_{2i}} + \frac{\mathbf{u}_{1i}^T \boldsymbol{\omega}^* \mathbf{u}_{2i}}{\lambda_i \mu_i \mathbf{v}_{1i}^T \boldsymbol{\omega}^* \mathbf{v}_{2i}} \right)^2 + \left( \frac{\mathbf{u}_{2i}^T \boldsymbol{\omega}^* \mathbf{u}_{2i}}{\lambda_i^2 \mathbf{v}_{1i}^T \boldsymbol{\omega}^* \mathbf{v}_{1i}} + \frac{\mathbf{u}_{1i}^T \boldsymbol{\omega}^* \mathbf{u}_{2i}}{\lambda_i \mu_i \mathbf{v}_{1i}^T \boldsymbol{\omega}^* \mathbf{v}_{2i}} \right)^2 \right) \quad (6.9)$$

Dabei beschreibt der Gewichtungsfaktor  $w_i$  das Vertrauen in die  $i$ -te Fundamental-Matrix  $\mathbf{F}_i$ . Die Zuverlässigkeit kann z.B. durch die Anzahl der verwendeten homologen Punkte abgeschätzt werden.

Zum Schluss kann die gesuchte Kalibrierungs-Matrix  $\mathbf{K}$  mit Hilfe einer CHOLESKY-Zerlegung aus der KRUPPA-Matrix  $\boldsymbol{\omega}^* = \mathbf{K}\mathbf{K}^T$  extrahiert werden. Dafür eignet sich z.B. der effiziente Algorithmus CHOLDC [PRE02].

### 6.3.2 Identische Singulärwerte der essentiellen Matrix

Von MENDONCA und CIPOLLA [MEN99] wurde eine einfache Methode zur Auto-Kalibrierung mit Hilfe der essentiellen Matrix vorgeschlagen. Dabei wird die Eigenschaft ausgenutzt, dass die beiden Singulärwerte identisch sind (vgl. Abschnitt 4.3.7).

Mit einer zunächst geschätzten Kalibrierungs-Matrix  $\mathbf{K}$  und der Annahme einer konstanten inneren Orientierung lässt sich für jede Bildkombination die essentielle Matrix  $\mathbf{E}$  aus der berechneten Fundamental-Matrix  $\mathbf{F}$  bestimmen

$$\mathbf{E}_i = \mathbf{K}^T \mathbf{F}_i \mathbf{K}. \quad (6.10)$$

Unter Verwendung der beiden sortierten Singulärwerte (vgl. Anhang A.1)

$$\mathbf{E} = \mathbf{U} \text{diag}(\lambda, \mu, 0) \mathbf{V}^T, \quad \lambda \geq \mu \quad (6.11)$$

kann für  $n$  Bilder die folgende Fehlerfunktion

$$\min_{\mathbf{K}} \sum_{i=1}^{\binom{n}{2}} w_i \left( 1 - \frac{\mu_i}{\lambda_i} \right) \quad (6.12)$$

aufgestellt werden. Die fünf gesuchten Parameter der Kalibrierungs-Matrix  $\mathbf{K}$  werden dabei so optimiert, dass für jede essentielle Matrix die beiden Singulärwerte möglichst identisch werden. Der Gewichtungsfaktor  $w_i$  beschreibt auch hier das Vertrauen in die entsprechende Fundamental-Matrix  $\mathbf{F}_i$ .

## 6.4 Kalibrierung mit dem dualen absoluten Quadric

### 6.4.1 Lineare Auto-Kalibrierung

Die Auto-Kalibrierung mit dem dualen absoluten Quadric  $\mathbf{Q}_\infty^*$  ist ein lineares Verfahren der Klasse B und erlaubt im Gegensatz zu den klassischen Methoden auch variierende Parameter für die innere Orientierung. Da anstelle der Fundamental-Matrizen die projektiven Rekonstruktionen verwendet werden, muss als Voraussetzung die projektive Basis mit den Verfahren aus Abschnitt 5.2 angepasst werden.

Anschließend kann mit

$$\boldsymbol{\omega}_i^* = \mathbf{K}_i \mathbf{K}_i^T = \lambda_i \begin{bmatrix} \alpha_x^2 + s^2 + x_0^2 & s\alpha_y + x_0 y_0 & x_0 \\ s\alpha_y + x_0 y_0 & \alpha_y^2 + y_0^2 & y_0 \\ x_0 & y_0 & 1 \end{bmatrix} = \mathbf{P}_i \mathbf{Q} \mathbf{Q}^T \mathbf{P}_i^T = \mathbf{P}_i \mathbf{Q}_\infty^* \mathbf{P}_i^T \quad (6.13)$$

ein Zusammenhang zwischen den berechneten Projektions-Matrizen  $\mathbf{P}_i$  und den gesuchten Parametern von  $\mathbf{Q}_\infty^*$  und  $\boldsymbol{\omega}_i^*$  aufgestellt werden [POL99b]. Dieser allgemeine Ansatz ist linear, wenn der Bildhauptpunkt im Ursprung angenommen wird (vgl. Tab. 6.1).

Um solche Bedingungen für die Parameter der inneren Orientierung in den Elementen von  $\boldsymbol{\omega}^*$  zu definieren, kann die Gleichung

$$\boldsymbol{\omega}_{mn}^* = \left( \mathbf{P} \mathbf{Q}_\infty^* \mathbf{P}^T \right)_{mn} = \mathbf{p}_m^T \mathbf{Q}_\infty^* \mathbf{p}_n \quad \text{mit} \quad \mathbf{P} = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3]^T \quad (6.14)$$

in der entsprechenden Zeile  $m$  und Spalte  $n$  angewendet werden.

Unter der Annahme, dass die Scherung vernachlässigt werden kann ( $\boldsymbol{\omega}_{12}^* = 0$ ) und sich der Bildhauptpunkt im Ursprung befindet ( $\boldsymbol{\omega}_{13}^* = \boldsymbol{\omega}_{23}^* = 0$ ), ergeben sich die homogenen Gleichungen

$$\mathbf{p}_1^T \mathbf{Q}_\infty^* \mathbf{p}_2 = \mathbf{p}_1^T \mathbf{Q}_\infty^* \mathbf{p}_3 = \mathbf{p}_2^T \mathbf{Q}_\infty^* \mathbf{p}_3 = 0. \quad (6.15)$$

Der geometrische Einfluss dieser drei *Orthogonalitäts-Bedingungen* auf die Abbildungsgeometrie der Projektions-Matrix  $\mathbf{P}$  ist in Abb. 6.7 dargestellt (vgl. Abschnitt 4.2.3.1).

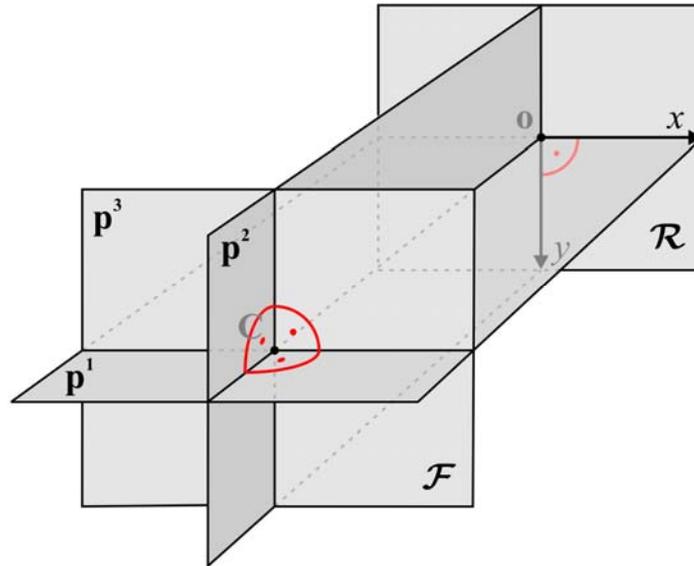
Optional kann auch ein bekanntes Seitenverhältnis  $\gamma$  mit

$$\gamma \mathbf{p}_1^T \mathbf{Q}_\infty^* \mathbf{p}_1 = \mathbf{p}_2^T \mathbf{Q}_\infty^* \mathbf{p}_2 \quad (6.16)$$

als Bedingung eingeführt werden. Dadurch vereinfacht sich die Gleichung (6.13) zu

$$\boldsymbol{\omega}_i^* = \lambda_i \begin{bmatrix} \alpha_i^2 & 0 & 0 \\ 0 & \alpha_i^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{P}_i \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \\ q_1 & q_4 & q_5 & q_6 \\ q_2 & q_5 & q_7 & q_8 \\ q_3 & q_6 & q_8 & q_9 \end{bmatrix} \mathbf{P}_i^T, \quad (6.17)$$

wobei  $\alpha_i$  der Kamerakonstanten im Bild  $i$  entspricht.



**Abb. 6.7:** Orthogonalitäts-Bedingungen für die Projektions-Matrix  $\mathbf{P}$

Um die zehn unbekannt Parameter  $q_k$  der symmetrischen Matrix  $\mathbf{Q}_\infty^*$  zu berechnen (vgl. Abschnitt 4.1.4.5), kann durch Umformen der Orthogonalitäts-Bedingungen (6.15) für  $n \geq 4$  Bilder ein lineares homogenes Gleichungssystem der Form

$$\mathbf{A}\mathbf{q} = \mathbf{0} \quad \text{mit} \quad \mathbf{q} = (q_0, q_1, \dots, q_9)^T \quad (6.18)$$

erstellt und durch eine Singulärwert-Zerlegung gelöst werden (vgl. Anhang A.2.1). Dabei liefert jede Projektions-Matrix  $\mathbf{P}_i$  folgende drei Zeilen der Design-Matrix:

$$\mathbf{A}_i = \begin{bmatrix} p_{11}p_{21} & p_{11}p_{31} & p_{21}p_{31} \\ p_{11}p_{22} + p_{12}p_{21} & p_{11}p_{32} + p_{12}p_{31} & p_{21}p_{32} + p_{22}p_{31} \\ p_{11}p_{23} + p_{13}p_{21} & p_{11}p_{33} + p_{13}p_{31} & p_{21}p_{33} + p_{23}p_{31} \\ p_{11}p_{24} + p_{14}p_{21} & p_{11}p_{34} + p_{14}p_{31} & p_{21}p_{34} + p_{24}p_{31} \\ p_{12}p_{22} & p_{12}p_{32} & p_{22}p_{32} \\ p_{12}p_{23} + p_{13}p_{22} & p_{12}p_{33} + p_{13}p_{32} & p_{22}p_{33} + p_{23}p_{32} \\ p_{12}p_{24} + p_{14}p_{22} & p_{12}p_{34} + p_{14}p_{32} & p_{22}p_{34} + p_{24}p_{32} \\ p_{13}p_{23} & p_{13}p_{33} & p_{23}p_{33} \\ p_{13}p_{24} + p_{14}p_{23} & p_{13}p_{34} + p_{14}p_{33} & p_{23}p_{34} + p_{24}p_{33} \\ p_{14}p_{24} & p_{14}p_{34} & p_{24}p_{34} \end{bmatrix}^T \quad (6.19)$$

Wird zusätzlich mit der Gleichung (6.16) die Bedingung für das bekannte Seitenverhältnis  $\gamma$  verwendet und  $\mathbf{A}$  um die Zeile

$$\mathbf{A}_i = \begin{bmatrix} \gamma p_{11}^2 - p_{21}^2 \\ 2(\gamma p_{11}p_{12} - p_{21}p_{22}) \\ 2(\gamma p_{11}p_{13} - p_{21}p_{23}) \\ 2(\gamma p_{11}p_{14} - p_{21}p_{24}) \\ \gamma p_{12}^2 - p_{22}^2 \\ 2(\gamma p_{12}p_{13} - p_{22}p_{23}) \\ 2(\gamma p_{12}p_{14} - p_{22}p_{24}) \\ \gamma p_{13}^2 - p_{23}^2 \\ 2(\gamma p_{13}p_{14} - p_{23}p_{24}) \\ \gamma p_{14}^2 - p_{24}^2 \end{bmatrix}^T \quad (6.20)$$

ergänzt, reduziert sich die Anzahl der benötigten Bilder auf drei (vgl. Tab. 6.1).

Nach der Berechnung des dualen absoluten Quadric  $\mathbf{Q}_\infty^*$  und mit der Kenntnis seiner Form im metrischen Raum  $\mathbf{Q}_M^* = \text{diag}(1,1,1,0)$  (vgl. Abschnitt 4.1.4.5) kann die unbekannte Transformation  $\mathbf{T}$  vom metrischen zum projektiven Raum bestimmt werden

$$\mathbf{Q}_\infty^* = \mathbf{T}\mathbf{Q}_M^*\mathbf{T}^T = \mathbf{T}\text{diag}(1,1,1,0)\mathbf{T}^T. \quad (6.21)$$

Dazu muss eine Eigensystem-Zerlegung  $\mathbf{Q}_\infty^* = \mathbf{E}\mathbf{D}\mathbf{E}^T$  (z.B. mit JACOBI [PRE02]) durchgeführt werden und durch Anpassen der sortierten Diagonalelemente von  $\mathbf{D}$

$$\mathbf{T} = \mathbf{E}\text{diag}(\sqrt{d_1}, \sqrt{d_2}, \sqrt{d_3}, 0) \quad \text{für} \quad d_1 \geq d_2 \geq d_3 \quad (6.22)$$

lässt sich die gesuchte Transformation bestimmen. Um zu verhindern, dass die Matrix  $\mathbf{T}$  genau wie  $\mathbf{Q}_\infty^*$  singulär wird, muss anschließend noch das Element  $t_{44} = 1$  gesetzt werden. Zum Abschluss können die projektiv rekonstruierten Objektpunkte und Projektionsmatrizen mit der inversen Homographie

$$\mathbf{X}_M = \mathbf{T}^{-1}\mathbf{X}_p \quad \text{bzw.} \quad \mathbf{P}_M = \mathbf{P}_p\mathbf{T} \quad (6.23)$$

in den metrischen Raum transformiert werden.

## 6.4.2 Verbesserungen und Erweiterungen

Für eine stabile Lösung von linearen homogenen Gleichungssystemen ist eine entsprechende Normierung der Daten erforderlich (vgl. Anhang A.2.1). Da ein geeigneter Normierungsfaktor a priori nicht bekannt ist, können die Elemente der Design-Matrix  $\mathbf{A}$  aus Gleichung (6.18) in einem weiteren Durchgang mit

$$\lambda = \frac{1}{\mathbf{p}_3^T \mathbf{Q}_\infty^* \mathbf{p}_3} \quad (6.24)$$

gewichtet werden, so dass nach wenigen Iterationen  $\omega_{33}^* = q_7 = 1$  gilt.

In POLLEFEYS u.a. [POL02] werden unterschiedliche Konstanten zur Normierung des linearen Gleichungssystems vorgeschlagen. Für die Parameter der inneren Orientierung ergeben sich dann z.B. folgende Gewichtungen

$$\begin{aligned} \frac{1}{9}(\mathbf{p}_1^T \mathbf{Q}_\infty^* \mathbf{p}_1 - \mathbf{p}_3^T \mathbf{Q}_\infty^* \mathbf{p}_3) &= 0, & 10(\mathbf{p}_1^T \mathbf{Q}_\infty^* \mathbf{p}_2) &= 0, \\ \frac{1}{9}(\mathbf{p}_2^T \mathbf{Q}_\infty^* \mathbf{p}_2 - \mathbf{p}_3^T \mathbf{Q}_\infty^* \mathbf{p}_3) &= 0, & 10(\mathbf{p}_1^T \mathbf{Q}_\infty^* \mathbf{p}_3) &= 0, \\ 5(\mathbf{p}_1^T \mathbf{Q}_\infty^* \mathbf{p}_1 - \mathbf{p}_2^T \mathbf{Q}_\infty^* \mathbf{p}_2) &= 0 \quad \text{und} \quad 100(\mathbf{p}_2^T \mathbf{Q}_\infty^* \mathbf{p}_3) &= 0, \end{aligned} \quad (6.25)$$

die aus den in der Praxis üblichen Wertebereichen hergeleitet wurden. Auch hier kann die initiale Lösung mit der Gleichung (6.24) noch iterativ verbessert werden.

Bei stark verrauschten Messwerten ist es möglich, dass  $\mathbf{Q}_\infty^*$  nicht wie gefordert positiv-definiert ist. Um dennoch eine Eigensystem-Zerlegung durchführen zu können, ist von SEO und HONG [SEO01] ein komplexwertiger Ansatz entwickelt worden.

### 6.4.3 Nicht-lineare Optimierung

Ist durch das lineare Verfahren in Abschnitt 6.4.1 erst einmal eine initiale Schätzung für den dualen absoluten Quadric  $\mathbf{Q}_\infty^*$  vorhanden, so kann diese mit einer nicht-linearen Optimierung verbessert werden. Um die beiden Seiten der Gleichung (6.13) möglichst gut anzupassen, eignet sich die Fehlerfunktion

$$\min_{\mathbf{K}_i, \mathbf{Q}_\infty^*} \sum_{i=1}^n \left\| N(\mathbf{K}_i \mathbf{K}_i^T) - N(\mathbf{P}_i \mathbf{Q}_\infty^* \mathbf{P}_i^T) \right\|, \quad (6.26)$$

wobei die Differenz der beiden Matrizen mit der Frobenius-Norm bestimmt wird und

$$N(\mathbf{A}) = \frac{\mathbf{A}}{\|\mathbf{A}\|} \quad (6.27)$$

eine Normierung der Matrix  $\mathbf{A}$  bezeichnet.

Wenn als erste Kamera  $\mathbf{P}_1 = [\mathbf{I} | \mathbf{0}]$  angenommen werden kann, so reduziert die *minimale Parametrisierung* von  $\mathbf{Q}_\infty^*$  mit der idealen Ebene  $\mathbf{\Pi}_\infty = (\mathbf{q}^T, 1)^T = (q_1, q_2, q_3, 1)^T$  die Anzahl der Variablen

$$\begin{aligned} \mathbf{Q}_\infty^* &= \begin{bmatrix} \boldsymbol{\omega}_1^* & -\boldsymbol{\omega}_1^* \mathbf{q} \\ -\mathbf{q}^T \boldsymbol{\omega}_1^* & \mathbf{q}^T \boldsymbol{\omega}_1^* \mathbf{q} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_1 \mathbf{K}_1^T & -\mathbf{K}_1 \mathbf{K}_1^T \mathbf{q} \\ -\mathbf{q}^T \mathbf{K}_1 \mathbf{K}_1^T & \mathbf{q}^T \mathbf{K}_1 \mathbf{K}_1^T \mathbf{q} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1^2 & 0 & 0 & -\alpha_1^2 q_1 \\ 0 & \alpha_1^2 & 0 & -\alpha_1^2 q_2 \\ 0 & 0 & 1 & -q_3 \\ -\alpha_1^2 q_1 & -\alpha_1^2 q_2 & -q_3 & \alpha_1^2 (q_1^2 + q_2^2) + q_3^2 \end{bmatrix}. \end{aligned} \quad (6.28)$$

Die nicht-lineare Optimierung des dualen absoluten Quadric vereinfacht sich dann zu

$$\min_{\alpha_i, \mathbf{\Pi}_\infty} \sum_{i=2}^n \left\| N(\mathbf{K}_i \mathbf{K}_i^T) - N(\mathbf{P}_i \mathbf{Q}_\infty^* \mathbf{P}_i^T) \right\| \quad (6.29)$$

mit nur  $3 + (n - 1)$  Parametern.

#### 6.4.4 Kombiniertes Verfahren

Eine weitere Verbesserung kann dadurch erreicht werden, dass neben der nicht-linearen Fehlerfunktion (6.29) gleichzeitig auch die Differenzen von den Singulärwerten der essentiellen Matrizen minimiert werden (vgl. Abschnitt 6.3.2). Dazu wird die folgende Kombination vorgeschlagen:

$$\min_{\alpha_i, \mathbf{\Pi}_\infty} \sum_{i=2}^n C_1(\alpha_i, \mathbf{\Pi}_\infty) + w \cdot C_2(\alpha_i) \quad \text{mit} \quad (6.30)$$

$$C_1(\alpha_i, \mathbf{\Pi}_\infty) = \left\| N(\mathbf{K}_i \mathbf{K}_i^T) - N(\mathbf{P}_i \mathbf{Q}_\infty^* \mathbf{P}_i^T) \right\| \quad \text{und} \quad C_2(\alpha_i) = \left( 1 - \frac{\mu_i}{\lambda_i} \right),$$

wobei  $w$  einen optionalen Gewichtungsfaktor darstellt.

Zunächst muss für eine Projektions-Matrix  $\mathbf{P}_i$  mit der Gleichung (4.66) die Fundamental-Matrix  $\mathbf{F}$  zur ersten Projektions-Matrix  $\mathbf{P}_1$  berechnet werden. Dann ist die essentielle Matrix  $\mathbf{E}$  mit der geschätzten Kalibrierungs-Matrix  $\mathbf{K}_i$  durch Anwenden der Gleichung (4.71) gegeben. Schließlich können die beiden Singulärwerte  $\lambda$  und  $\mu$  durch eine SVD bestimmt werden (vgl. Anhang A.1).

Der Vorteil dieser neuen Kombination ist in Tab. 6.2 am Beispiel der VIKTORIA-Statue dargestellt. Bei der Berechnung wurde der Gewichtungsfaktor  $w = 0,01$  verwendet. Mit diesem empirisch bestimmten Wert wurden in der Praxis die besten Resultate erzielt.

### 6.5 Experimentelle Ergebnisse

Um die Leistungsfähigkeit der vorgestellten Methoden zur Selbst-Kalibrierung einschätzen zu können, sind 25 Bilder der NOFRETETE-Büste mit  $1.020 \times 1.360$  Pixeln und 19 Bilder der VIKTORIA-Statue mit  $768 \times 1.024$  Pixeln automatisch ausgewertet worden.

Die Ergebnisse bei der Rekonstruktion der *äußeren Orientierung* sind in Abb. 5.7 und Abb. 6.8 dargestellt. Die entsprechenden Parameter der *inneren Orientierung* sind in Tab. 6.3 aufgeführt. Für die Bewertung der Resultate wurden teilweise Passpunkte und elektronisch aufgezeichnete Kalibrierungsinformationen verwendet.



**Abb. 6.8:** Rekonstruktion der äußeren Orientierung am Beispiel der VIKTORIA-Statue.  
Die Kameras sind zur Illustration manuell in einen Luftbildausschnitt der HRSC-A vom DLR Berlin eingepasst worden.

Ansatz	$C_1$	$C_2$	Fehler von $\alpha$ [in Prozent]		CPU [in Sek.]
			$\emptyset$	Max	
Linear	0,044	0,218	8,54	14,87	0,01
Nicht-linear	0,006	0,151	4,74	12,72	0,39
Kombiniertes Verfahren	0,007	0,003	2,99	9,12	1,15

**Tab. 6.2:** Vergleich der Methoden zur Auto-Kalibrierung mit dem dualen absoluten Quadric bei bekannten Kamerakonstanten  $\alpha$

Kamerakonstante						Bildhauptpunkt [in Pixel]				Seitenverhältnis $\gamma$		Scherung $\theta$ [in Grad]	
$\alpha$ [in Pixel]			$C$ [in mm]			$x_0$		$y_0$					
Soll	$\emptyset$	Max	Soll	$\emptyset$	Max	$\emptyset$	Max	$\emptyset$	Max	$\emptyset$	Max	$\emptyset$	Max
<b>NOFRETETE-Büste (SONY DCR-TRV 30E)</b>													
3.639	120	463	34	1,12	4,33	75	131	57	178	0,001	0,004	0,11	0,40
<b>VIKTORIA-Statue (CANON EOS-IX)</b>													
6.179-6.477	194	591	311-326	9,78	29,78	47	86	13	23	0,012	0,034	0,02	0,04

**Tab. 6.3:** Abweichung der automatisch kalibrierten Parameter für die innere Orientierung von den tatsächlich gemessenen Werten

## 6.6 Diskussion

In diesem Kapitel wurde das Prinzip der automatischen Selbst-Kalibrierung aus mindestens drei Bildern erläutert. Der lineare Ansatz mit dem dualen absoluten Quadric erlaubt die Berechnung von variierenden Parametern bei fokussierbaren Kameras im Nahbereich. Außer den drei Orthogonalitäts-Bedingungen für die innere Orientierung werden keine weiteren Informationen benötigt. Sollten dennoch zusätzliche Kalibrierdaten bekannt sein, so können diese flexibel in das Verfahren integriert werden. Für die nicht-lineare Optimierung des dualen absoluten Quadric wurde ein neues Kriterium vorgeschlagen, das die Singulärwerte der essentiellen Matrix berücksichtigt.

Die bei der automatischen Selbst-Kalibrierung erzielte Genauigkeit ist nicht mit denen von Labor-Kalibrierungen moderner Messkameras zu vergleichen. Gerade solche Parameter, die einen geringen Einfluss auf die 3D-Rekonstruktion haben, werden häufig nicht besonders präzise bestimmt. Dennoch sind die so ermittelten Orientierungsdaten für eine photogrammetrische Nutzung der Bilder geeignet.

*„Ein Fachmann ist ein Mensch,  
der zu reden anfängt,  
wenn er zu denken aufhört  
– und umgekehrt.“*

UMBERTO ECO [\* 1932];  
ital. Schriftsteller

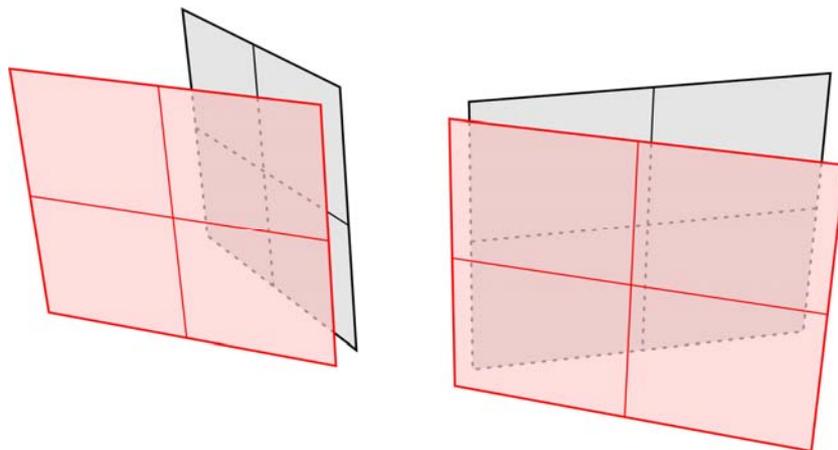
## 7 Metrische 3D-Rekonstruktion

Nachdem die Aufnahmegeometrie aus wenigen Punktmerkmalen bestimmt wurde, kann anschließend die Form des Objektes rekonstruiert werden. Die Erstellung von *Oberflächenmodellen* erfolgt üblicherweise durch eine flächenhafte Korrespondenz-Analyse. Die aufwändige Berechnung kann allerdings erheblich vereinfacht und beschleunigt werden, wenn die konvergenten Bildpaare in den Stereo-Normalfall transformiert werden (vgl. Abschnitt 7.1). Um dabei kritische Konfigurationen zu vermeiden, wird ein hybrider Ansatz beschrieben, der durch die Verwendung von Polarkoordinaten beliebige Bildanordnungen im Nahbereich erlaubt.

Abschließend wird als Anwendungsbeispiel eine *volumetrische 3D-Rekonstruktion* vorgestellt. Dabei wird gezeigt, wie durch Ausschneiden von Objektkonturen aus einem Volumen relativ einfach ein 3D-Modell erstellt werden kann (vgl. Abschnitt 7.2).

### 7.1 Normalbild-Erstellung

Konvergente Stereobildpaare können so umgerechnet werden, dass sie dem Stereo-Normalfall entsprechen (vgl. Abb. 7.1). Da die Epipolar-Geraden parallel zur x-Richtung liegen, vereinfacht sich die Korrespondenz-Analyse zu einer eindimensionalen horizontalen Suche. Die Normalbild-Erstellung wird in der Regel durch Anwenden einer linearen planaren Transformation erreicht, wie z.B. von HARTLEY [HAR99] beschrieben. Dieses Verfahren ist relativ einfach, schnell und bewahrt Bildmerkmale, wie z.B. gerade Linien.



**Abb. 7.1:** Geometrie von Normalbildern

Der lineare Ansatz ist allerdings nicht allgemein und schlägt fehl, wenn sich der Epipol innerhalb des Bildes befindet. Die Transformation mit einem Epipol, der sich dicht am Bildrand befindet, führt zu einem extrem großen und stark verzerrten Bild. Durch die Verwendung von Stereo-Anordnungen mit nahezu parallel ausgerichteten Kameras lässt sich dieses Problem vermeiden. Bei allgemeinen Bildsequenzen mit einer beliebigen Kameraausrichtung muss für die Normalbild-Erstellung jedoch eine verbesserte Methode eingesetzt werden.

Um dieses Problem zu lösen, schlugen ROY, MEUNIER und COX [ROY97] eine zylindrische Entzerrungs-Methode mit einer separaten Transformation für jede Epipolar-Gerade vor. Der Grundgedanke liegt in der Verwendung von *Polarkoordinaten* mit dem *Epipol als Ursprung*. Allerdings wurden viele aufwändige Operationen im euklidischen Raum durchgeführt.

In den Arbeiten von POLLEFEYS [POL99a] wurde dieser nicht-lineare Ansatz verbessert und für die *projektive Geometrie* angepasst. Durch den Einsatz von unterschiedlichen nicht-linearen Transformationen für die beiden Bilder werden die Bildmerkmale allerdings unterschiedlich verzerrt, was eine anschließende Korrespondenz-Analyse erschwert. Um diesen Effekt zu vermeiden, soll ein *hybrides Verfahren* von ORAM [ORA01] vorgestellt werden, das zunächst die Epipole mit einer linearen Transformation in Übereinstimmung bringt (vgl. Abschnitt 7.1.1) und anschließend für beide Bilder die gleiche nicht-lineare Transformation verwendet (vgl. Abschnitt 7.1.2).

Das Verfahren benötigt ein Bildpaar mit einem gemeinsamen Überlappungsbereich und mindestens sieben homologen Bildpunkten  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ , so dass die Fundamental-Matrix  $\mathbf{F}$  und die Epipole  $\mathbf{e}$  bzw.  $\mathbf{e}'$  bestimmt werden können (vgl. Abschnitt 4.3.2). Als Ergebnis entsteht ein geometrisch transformiertes Bildpaar, bei dem die Epipolar-Geraden horizontal und alle verbleibenden Disparitäten parallel zur  $x$ -Achse verlaufen.

### 7.1.1 Lineare Bild-Transformation

In der ersten Stufe muss eines der beiden Bilder mit einer linearen Abbildung so transformiert werden, dass die Epipolar-Geraden der Bilder übereinstimmen. Für die Auswahl ist die Entfernung des jeweiligen Epipols vom Bildmittelpunkt entscheidend. Um das Problem mit extrem großen Ergebnisbildern zu reduzieren, wird das Bild mit dem vom Bildmittelpunkt weiter entfernten Epipol auf das entsprechende Referenzbild transformiert.

Für die planare Entzerrung kann eine Homographie eingesetzt werden (vgl. 4.1.5.2). Die zur *Epipolar-Geometrie kompatible Homographie*  $\mathbf{H}$  bildet jeden Punkt so ab, dass dieser auf der korrespondierenden Epipolar-Geraden liegt.

Nach der Anwendung der Homographie besitzen beide Bilder denselben Epipol

$$\mathbf{e}' = \mathbf{H}\mathbf{e} \quad \text{und} \quad \mathbf{e} = \mathbf{H}^{-1}\mathbf{e}'. \quad (7.3)$$

Eine Homographie ist mit der Fundamental-Matrix  $\mathbf{F}$  genau dann kompatibel, wenn die Matrix  $\mathbf{H}^T\mathbf{F}$  schiefssymmetrisch ist, d.h.

$$\mathbf{H}^T\mathbf{F} + \mathbf{F}^T\mathbf{H} = \mathbf{0} \quad (7.4)$$

gilt. Die Matrix  $\mathbf{H}$  ist folgendermaßen definiert:

$$\mathbf{H} = \mathbf{M} - \mathbf{e}'\mathbf{v}^T \quad \text{mit} \quad \mathbf{M} = [\mathbf{e}']_{\times} \mathbf{F}, \quad (7.5)$$

wobei der unbekannte 3-Vektor  $\mathbf{v}$  eine Ebene  $\Pi$  im projektiven Raum  $\mathbb{P}^3$  definiert, mit deren Hilfe die gesuchte Homographie  $\mathbf{H}$  Punkte zwischen den beiden Bildern transferriert (vgl. Abb. 7.2). Für die Berechnung von  $\mathbf{H}$  aus  $n \geq 3$  homologen Bildpunkten kann die Minimierung des folgenden Kriteriums betrachtet werden:

$$\min \sum_{i=1}^n d(\mathbf{x}'_i, \mathbf{H}\mathbf{x}_i)^2 = \min \sum_{i=1}^n d(\mathbf{x}'_i, (\mathbf{M} - \mathbf{e}'\mathbf{v}^T)\mathbf{x}_i)^2, \quad (7.6)$$

wobei  $d$  den (euklidischen) Abstand im Bild bezeichnet. Um den gesuchten Vektor  $\mathbf{v}$  zu berechnen, lassen sich durch Umformungen für jeden homologen Punkt zwei lineare Gleichungen aufstellen

$$(\mathbf{x}_i \mathbf{e}'_3 - \mathbf{e}'_i) \mathbf{x}'^T \mathbf{v} = (\mathbf{x}_i \mathbf{m}_3^T - \mathbf{m}_i) \mathbf{x}' \quad \text{für} \quad i = 1, 2, \quad (7.7)$$

wobei die Indizes die  $n$ -te Komponente des jeweiligen Vektors und  $\mathbf{m}_n$  die  $n$ -te Zeile von  $\mathbf{M}$  bezeichnen. Das Zusammenführen der  $2n$  Gleichungen ergibt ein lineares Gleichungssystem der Form  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , das z.B. mit einer inhomogenen linearen Kleinste-Quadrate-Lösung bestimmt werden kann (vgl. Anhang A.2.2). Dabei sollte auf den Sonderfall geachtet werden, dass die verwendeten Bildpunkte nicht kollinear angeordnet sind, da ansonsten keine Lösung existiert.

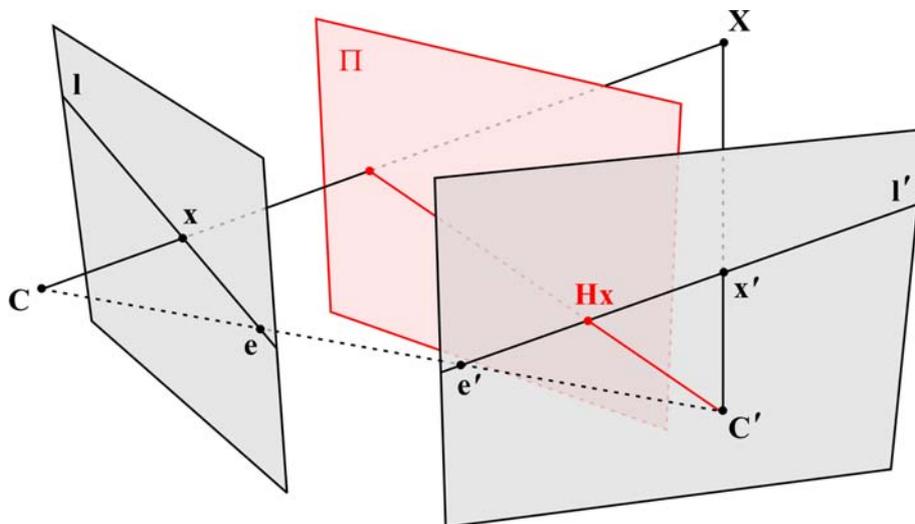
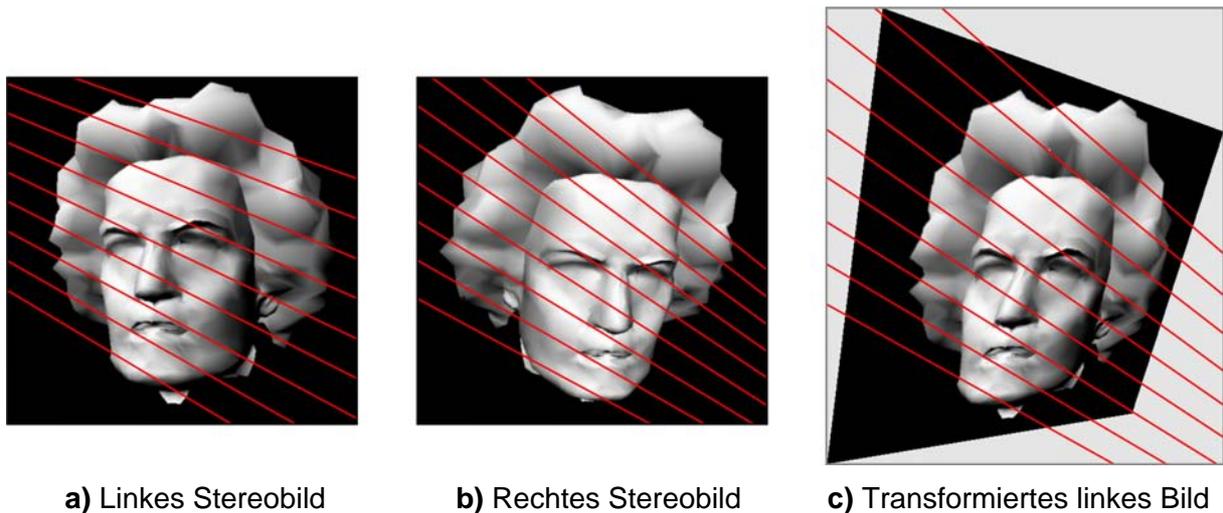


Abb. 7.2: Zur Epipolar-Geometrie kompatible Homographie



**Abb. 7.3:** Ergebnis der linearen Transformation zur Angleichung der Epipole

Da die gemessenen homologen Bildpunkte verrauscht sein können und nicht exakt der Epipolar-Geometrie entsprechen, sollten die Positionen vor der Berechnung korrigiert werden. Die verbesserten Bildpunkte, welche die Epipolar-Bedingung exakt erfüllen, können mit Hilfe der optimalen Triangulation aus Abschnitt 4.5.2 bestimmt werden.

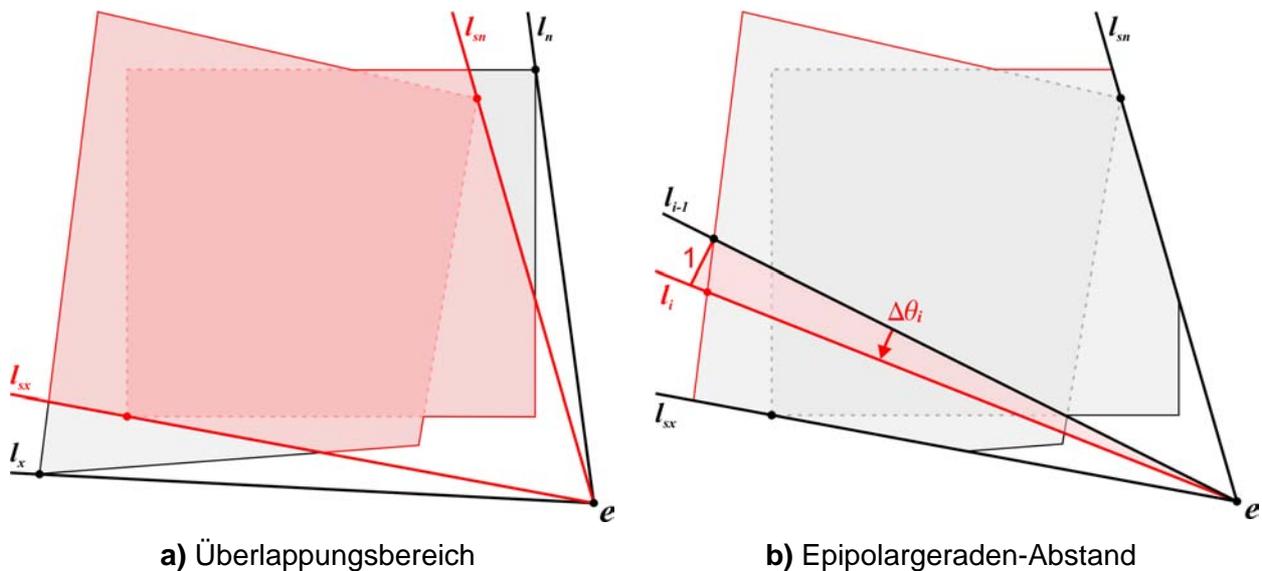
Abb. 7.3 zeigt die geometrische Transformation mit der zur Fundamental-Matrix  $\mathbf{F}$  kompatiblen Homographie  $\mathbf{H}$ . Das Bild wurde durch indirekte Entzerrung mit bikubischer Interpolation erzeugt (vgl. Abschnitt 3.1).

## 7.1.2 Nicht-lineare Bild-Transformation

Nachdem die Epipole mit Hilfe der Homographie in Übereinstimmung gebracht worden sind, müssen in einer zweiten Stufe die Epipolar-Geraden noch parallel zur  $x$ -Achse ausgerichtet werden. Dies wird dadurch erreicht, dass die Bildpunkte in Polarkoordinaten ausgedrückt werden, wobei der Epipol als Ursprung definiert wird. Um mit der Entzerrung fortzufahren zu können, sollten zunächst die Dimensionen der resultierenden Ergebnisbilder bestimmt werden.

### 7.1.2.1 Bestimmung der optimalen Bildgröße

Unter der Annahme, dass die Bilder rechteckig sind und sich der gemeinsame Epipol außerhalb der Bilder befindet, gehen die äußersten Epipolar-Geraden durch die Bildecken. Durch die Bestimmung der Polarwinkel für jede Ecke lässt sich ein minimaler und maximaler Winkel bestimmen, die maximal einen Bereich von  $\pi$  Radiant aufspannen. Diese Bereichseinschränkung ermöglicht eine für die Implementierung günstige Normierung der Winkel. Es empfiehlt sich, einen der Winkel auf 0 Radiant zu setzen und mit einfacher Minimum- und Maximumsuche fortzusetzen. Anschließend wird der Normierungswert wieder zu dem Ergebnis addiert.



**Abb. 7.4:** Bestimmung der optimalen Bildgröße

Wenn für beide Bilder die minimalen Winkel  $n$ ,  $n'$  und maximalen Winkel  $x$ ,  $x'$  berechnet wurden, ergibt sich der gesuchte Überlappungsbereich aus dem zweiten Minimum  $sn$  und dem zweiten Maximum  $sx$

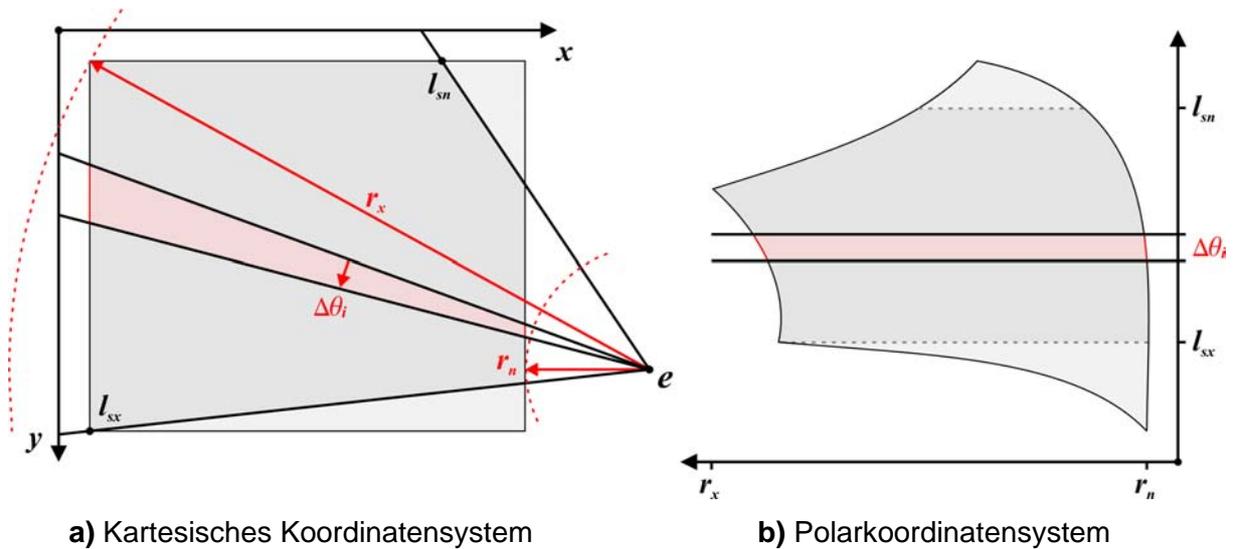
$$sn = \begin{cases} \min(n, n') & \|n - n'\| > \pi \\ \max(n, n') & \|n - n'\| \leq \pi \end{cases} \quad sx = \begin{cases} \max(x, x') & \|x - x'\| > \pi \\ \min(x, x') & \|x - x'\| \leq \pi \end{cases} \quad (7.8)$$

Die Bestimmung des Überlappungsbereichs ist in Abb. 7.4a skizziert.

Diese Vorgehensweise schlägt allerdings fehl, wenn sich der Epipol innerhalb des Bildes befindet, da der Winkel in diesem Fall  $2\pi$  Radiant umfasst. Wenn der Epipol nur in einem der beiden Bilder liegt, so können die gesuchten Winkel durch die Ecken des anderen Bildes bestimmt werden. Befindet sich der Epipol in beiden Bildern, so können der minimale und maximale Winkel einfach auf  $-\pi$  bzw.  $\pi$  gesetzt werden.

Damit keine Bildinformation verloren geht, sollte jedes Pixel mindestens einem Element in dem transformierten Bild entsprechen. Um einen möglichen Informationsverlust zu vermeiden, muss der kritische Bereich also mindestens ein Pixel umfassen. In der Regel befindet sich dieser an den Bildrändern mit der maximalen Entfernung zum Epipol.

Um gleichzeitig eine minimale Bildgröße zu gewährleisten, wird für jede Epipolar-Gerade eine optimale Breite berechnet (vgl. Abb. 7.4b). Angefangen von dem minimalen äußeren Begrenzungswinkel  $sn$  wird die nächste Epipolar-Gerade durch einen so kleinen Winkelschritt gefunden, dass der maximale Abstand benachbarter Epipolar-Geraden im Bild weniger als ein Pixel beträgt. Dieser kleine Winkelschritt  $\theta$  lässt sich leicht mit  $\theta = \arctan(1/d)$  bestimmen, wobei  $d$  durch Schneiden der Epipolar-Gerade mit dem entsprechenden Bildrand bestimmt werden kann.



**Abb. 7.5:** Transformation in Polardarstellung

Dieser Vorgang wird sukzessiv bis zum Erreichen des maximalen Begrenzungswinkels  $s_x$  wiederholt. Die verwendete Anzahl der Winkelschritte definiert die *Höhe* des Ergebnisbildes in Polardarstellung. Aufgrund des gemeinsamen Epipols ist die Zuordnung eines Winkels einer Epipolar-Gerade zu einer Zeilennummer in der Polardarstellung für beide Bilder identisch.

Um die Berechnung zu vereinfachen, empfiehlt sich für die Implementierung die Einführung einer Look-Up-Tabelle [ORA01]. Die *Breite* des Ergebnisbildes in Polardarstellung ergibt sich, indem für alle Epipolar-Geraden der minimale und maximale Abstand  $r_n$  bzw.  $r_x$  des Epipols vom Bildrand bestimmt wird (vgl. Abb. 7.5).

### 7.1.2.2 Transformation in Polarkoordinaten

Beide Bilder müssen nun vom  $(x, y)$ -Raum in den  $(r, \theta)$ -Raum transformiert werden. Aufgrund der indirekten geometrischen Transformation (vgl. Abschnitt 3.1) wird für eine ganzzahlige Polarkoordinate die Position des zu interpolierenden Intensitätswertes in den kartesischen Koordinaten gesucht

$$x = e_x + \cos(\theta) \cdot r \quad y = e_y + \sin(\theta) \cdot r \quad \text{für} \quad r_n \leq r \leq r_x \quad \text{und} \quad l_{sn} \leq \theta \leq l_{sx}. \quad (7.9)$$

Die Rücktransformation von kartesischen Koordinaten in die Polardarstellung ist entsprechend durch

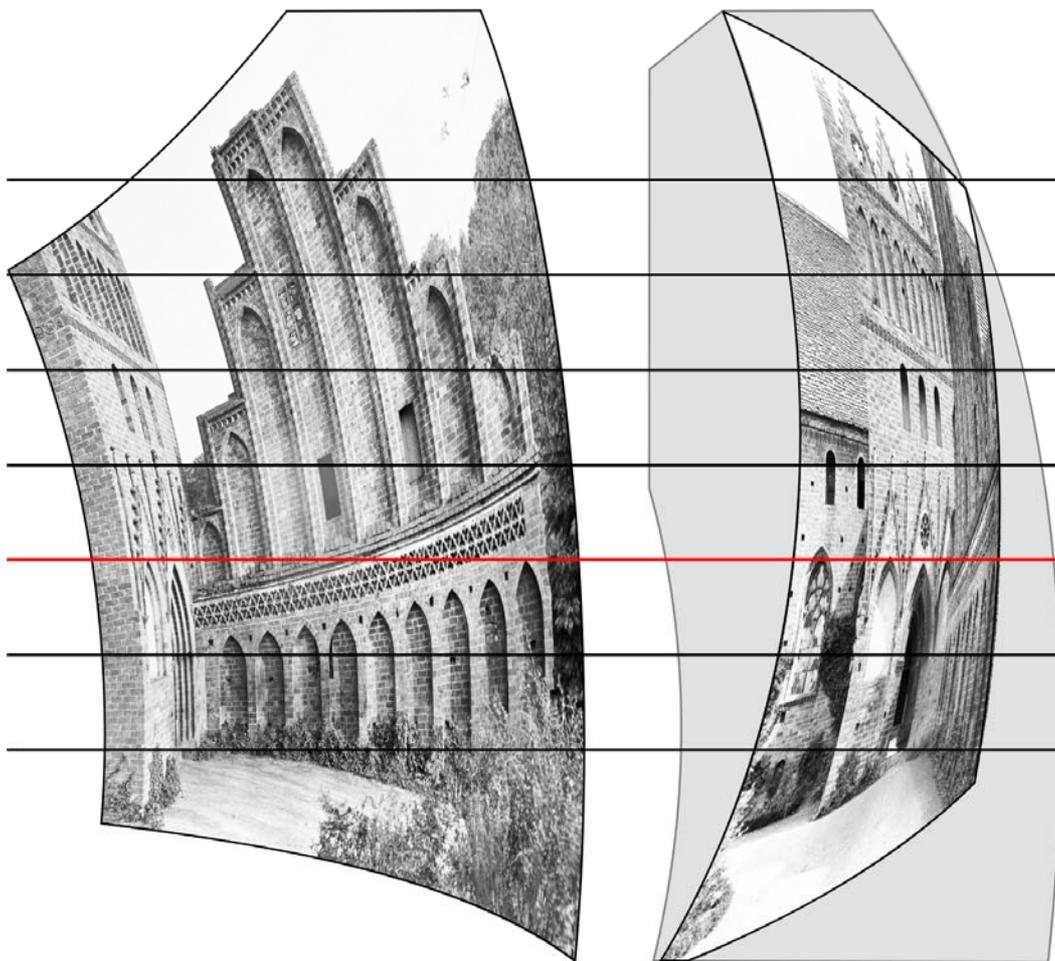
$$\theta = \arctan\left(\frac{y - e_y}{x - e_x}\right) \quad r = \frac{x - e_x}{\cos(\theta)} \quad \text{für} \quad -\pi \leq \theta \leq \pi \quad (7.10)$$

definiert.

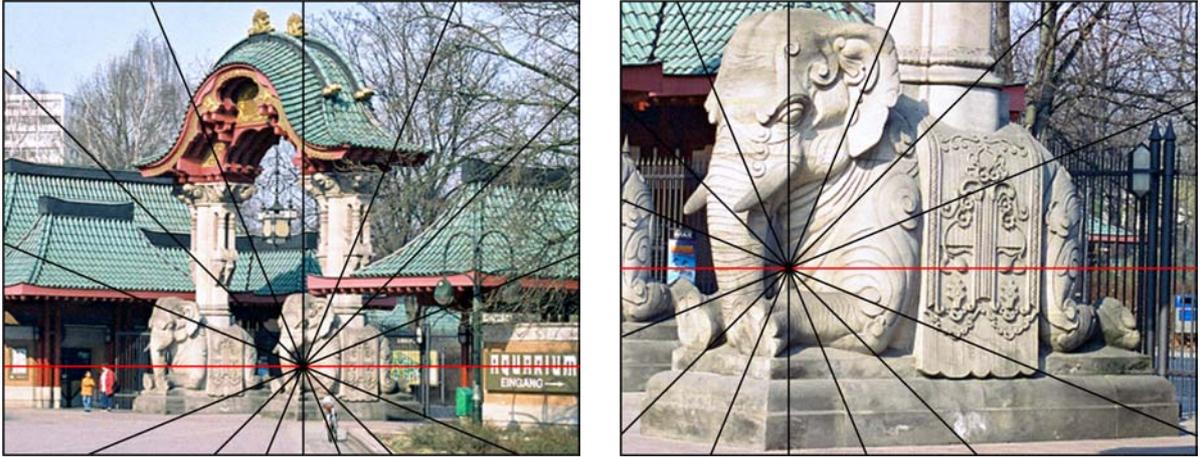
Die vorgestellte Methode funktioniert nicht bei Bildern mit unendlichen Epipolen, da alle Entfernungen ebenfalls unendlich und alle Winkel identisch werden. Dieser Sonderfall lässt sich allerdings relativ einfach beheben. Befindet sich in dem Referenzbild ein unendlicher Epipol, so brauchen nach der Anwendung der kompatiblen Homographie die beiden Bilder nur noch so gedreht werden, dass die Epipolar-Geraden parallel zur  $x$ -Achse liegen.

### 7.1.3 Experimentelle Ergebnisse

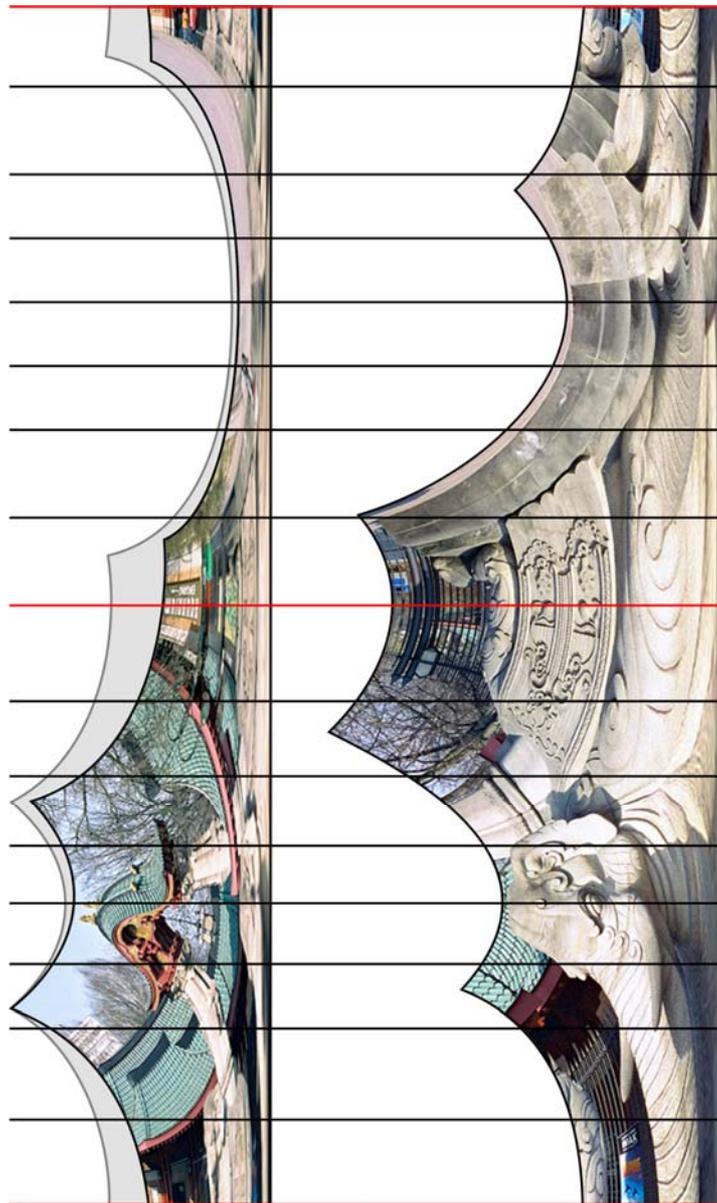
Um die Funktionsweise der Normalbild-Erstellung zu demonstrieren, wurden die Messbilder vom Kloster in CHORIN (vgl. Abb. 4.15) und ein Stereo-Bildpaar vom ELEFANTENTOR des Zoologischen Gartens in Berlin (vgl. Abb. 7.7) mit dem hybriden Verfahren in Polarkoordinaten transformiert.



**Abb. 7.6:** Normalbildpaar nach der Transformation in Polarkoordinaten



a) Bildpaar vom ELEFANTENTOR nach einer Vorwärtsbewegung der Kamera



b) Normalbilder in Polarkoordinaten

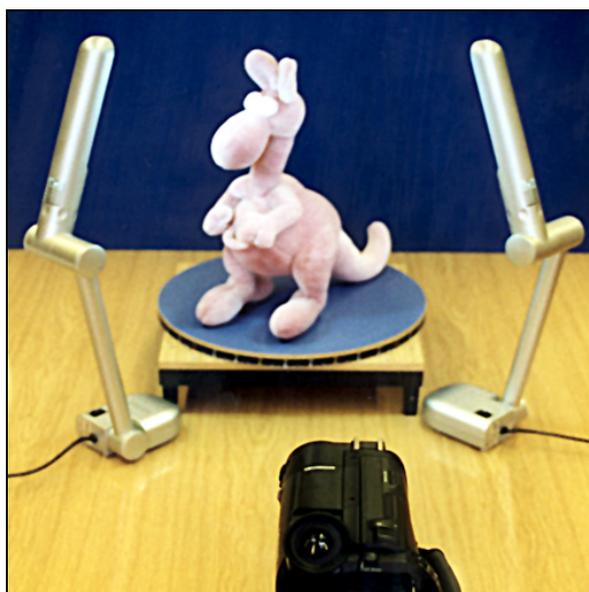
**Abb. 7.7:** Normalbild-Erstellung für ein Bildpaar mit abgebildeten Epipolen

## 7.2 Volumetrische 3D-Modellierung

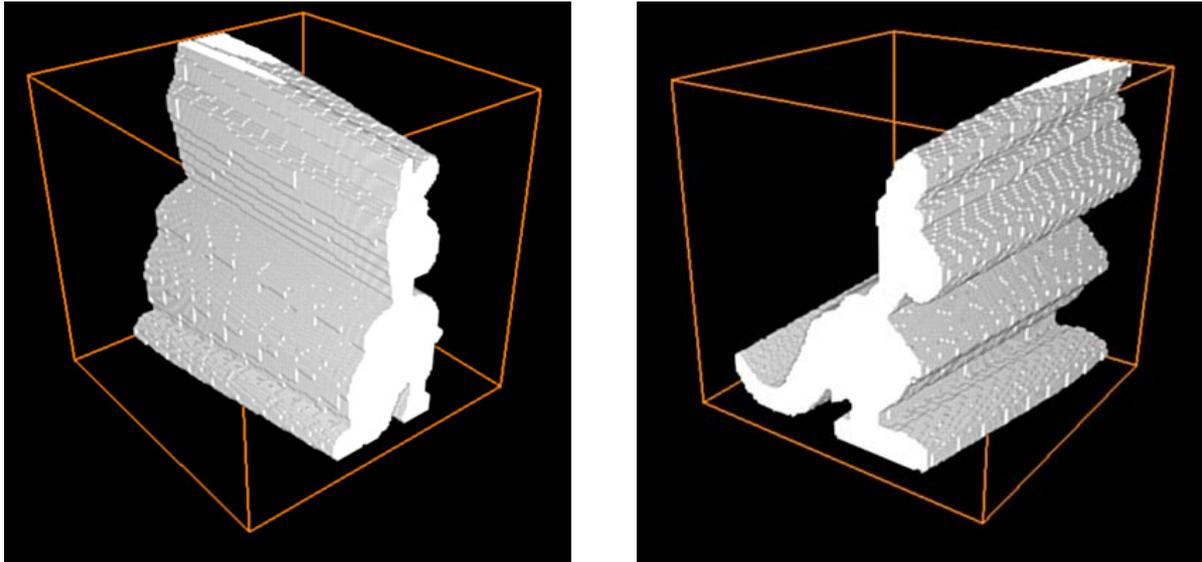
Die Form der abgebildeten Objekte kann durch eine flächenhafte Zuordnung der Normalbilder aus Abschnitt 7.1 automatisch rekonstruiert werden. Man erhält dadurch ein zusammenhängendes Oberflächenmodell, das eine so genannte *2,5D-Rekonstruktion* beschreibt. In der Photogrammetrie ist diese Auswertemethode weit entwickelt und in vieler Hinsicht ausgereift [HEI95].

Eine effiziente Methode zur Generierung einer vollständigen *3D-Rekonstruktion* besteht in der Berechnung eines Volumenmodells durch Auswerten von Objektkonturen (*Shape from Silhouette*, *Shape from Contour*) [KUZ02, FIT98, MAR83]. Dabei wird das Objekt häufig auf einen Drehteller gestellt, der sich vor einem definierten Hintergrund befindet (vgl. Abb. 7.8). Dadurch kann der Umriss der Objekte mit einem Segmentierungsverfahren automatisch bestimmt werden. Wenn die Abbildungsgeometrie bekannt ist, so lassen sich bei einer Zentralprojektion verschiedene Kegel von den Projektionszentren zu den Konturen als Basis aufspannen (vgl. Abb. 7.9).

Die Schnittmenge dieser Kegel definiert eine erste Schätzung der Objektform, die so genannte *visuelle Hülle*. Für die Berechnung muss jedes Volumenelement (*Voxel*)  $\mathbf{V} = (X, Y, Z, 1)^T$  mit allen euklidischen Projektions-Matrizen  $\mathbf{P}_i$  abgebildet werden. Das Voxel  $\mathbf{V}$  wird gelöscht, falls ein signifikanter Anteil der projizierten Bildpunkte außerhalb der segmentierten Objektkonturen liegt (*Voting-based carving*). Ein prinzipielles Problem bei diesem Ansatz stellen konkave Objektbereiche dar, wenn sie durch die Konturen verdeckt werden. Eine verbesserte Methode durch die Kombination mit einer flächenbasierten Bildzuordnung ist z.B. in KUZU und RODEHORST [KUZ01] beschrieben.



**Abb. 7.8:** Versuchsaufbau mit einem fixierten Camcorder und einem Drehteller vor einem definierten Hintergrund



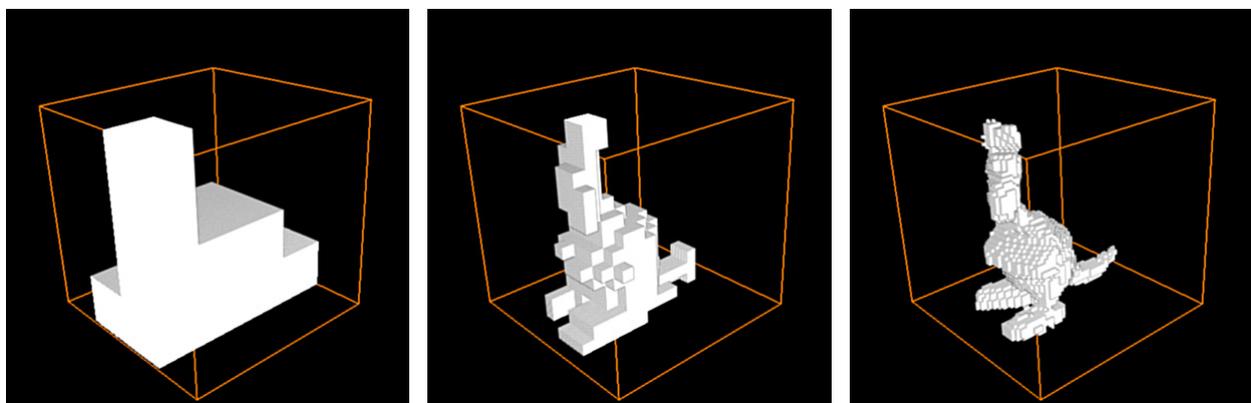
**Abb. 7.9:** Ausschneiden verschiedener Konturen unter Zentralprojektion

### 7.2.1 Hierarchische Berechnung mit Octrees

Die Anzahl der verwendeten Volumenelemente beeinflusst die Genauigkeit der 3D-Modellierung. In der Regel ist je nach verfügbarer Technik ein Kompromiss zu dem erforderlichen Speicher- und Berechnungsaufwand notwendig.

Durch das stufenweise Ausschließen von Bereichen, die keine relevante Objektinformation enthalten, kann der Auswerteprozess allerdings deutlich beschleunigt werden. Daher wurde eine hierarchische Berechnung mit einer einfachen *Octree-Struktur* realisiert [SRI90]. Die iterative Verfeinerung der Voxels ist in Abb. 7.10 veranschaulicht.

Der durch die hierarchische Prozessierung erreichte Geschwindigkeits-Vorteil ist in Abb. 7.11 dargestellt. Für das gezeigte Objekt wurde in fünf Auflösungsstufen die Berechnungszeit von ca. 12 Minuten auf 45 Sekunden reduziert, was einer Beschleunigung um den Faktor 16 entspricht.

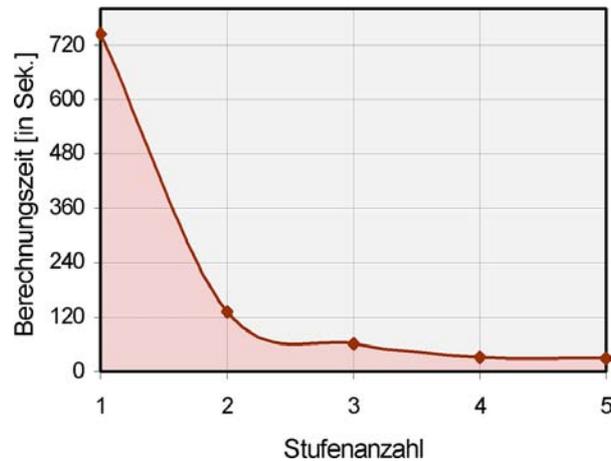


a)  $4^3$  Voxel

b)  $16^3$  Voxel

c)  $64^3$  Voxel

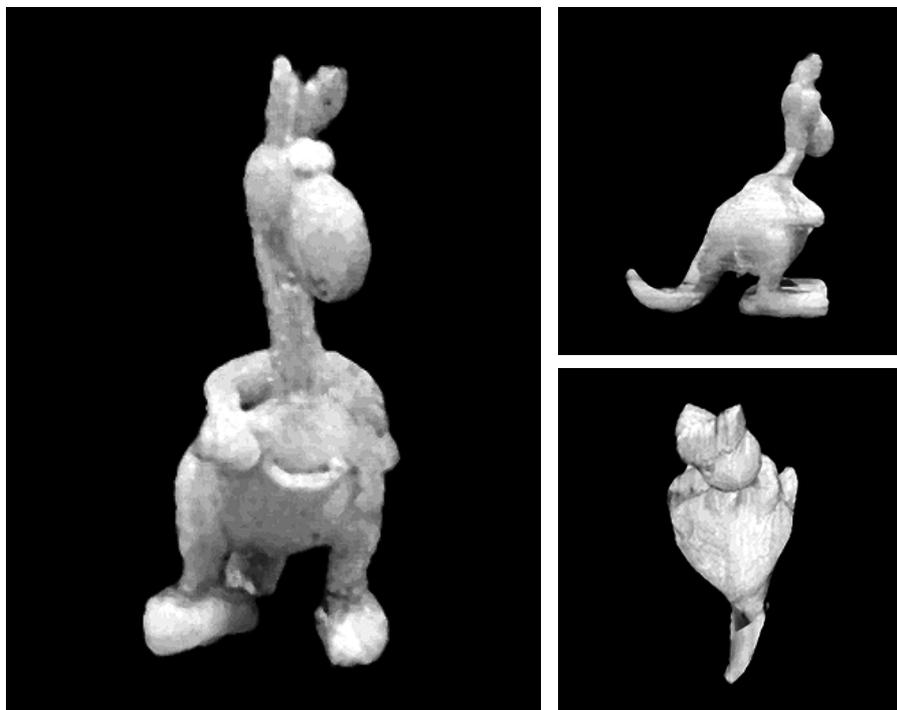
**Abb. 7.10:** Iterative Verfeinerung des Volumenmodells



**Abb. 7.11:** Berechnungszeit in Abhängigkeit von den verwendeten Auflösungsstufen

## 7.2.2 Experimentelle Ergebnisse

Abb. 7.12 zeigt das Ergebnis der volumetrischen 3D-Modellierung. Dabei wurden die Konturen von 30 Bildern automatisch extrahiert und mit Hilfe der rekonstruierten Orientierungsparameter aus einem Volumen von  $256^3$  Voxeln ausgeschnitten. Anschließend ist das Modell mit der abgebildeten Textur eingefärbt worden.



**Abb. 7.12:** Texturierte volumetrische 3D-Rekonstruktion

## 7.3 Diskussion

Es konnte gezeigt werden, wie mit Hilfe der projektiven relativen Orientierung die Umrechnung von konvergenten Bildpaaren in den Stereo-Normalfall realisiert werden kann. Die Erstellung von Oberflächenmodellen mit klassischen flächenbasierten Bildzuordnungs-Verfahren kann dadurch wesentlich einfacher und schneller erfolgen. Das hybride Verfahren erlaubt beliebige Aufnahmekonfigurationen im Nahbereich und erreicht eine optimale Bildgröße, indem nur der gemeinsame Überlappungsbereich berücksichtigt und für jede Epipolar-Gerade eine individuelle Breite berechnet wird.

Das Anwendungsbeispiel mit der volumetrischen 3D-Rekonstruktion für perspektivische Abbildungen ist ein einfaches, aber dennoch leistungsfähiges Verfahren, wenn die Objektkonturen in den Bildern erkannt werden können. Dabei wurde nachgewiesen, dass der erforderliche Berechnungsaufwand durch eine hierarchische Verfeinerung mit einer Octree-Struktur deutlich reduziert wird.

## 8 Zusammenfassung und Ausblick

Die vorliegende Arbeit soll dazu beitragen, dass die photogrammetrische 3D-Rekonstruktion im Nahbereich automatisiert und mit Hilfe einer Selbst-Kalibrierung flexibler gestaltet werden kann. Dazu sind Methoden für eine *merkmalsbasierte Korrespondenz-Analyse* und eine *Auto-Kalibrierung* mit Hilfe der projektiven Geometrie entwickelt und getestet worden. Durch die Integration und Verbesserung aktueller Forschungsergebnisse konnte ein durchgängiger Auswerteprozess vorgestellt werden. Das Potential und die Effizienz des innovativen Ansatzes wurden anhand verschiedener Anwendungen nachgewiesen.

Für die stabile Extraktion von signifikanten Punktmerkmalen wurden verschiedene *Interest-Operatoren* verglichen. Unter den Anforderungen der Bildzuordnung hat der FÖRSTNER-Operator die besten Ergebnisse erzielt. Die quantitative Analyse hat die Bedeutung einer robusten Gradientenberechnung unterstrichen. Für die Lokalisierung mit Subpixel-Genauigkeit, eine gleichmäßige Punktverteilung sowie die Nutzung von Farbinformation sind verschiedene Verbesserungen beschrieben und erfolgreich getestet worden. Allerdings sollte durch zukünftige Untersuchungen der beschriebene Lagefehler bei der Bestimmung von L- und T-Ecken reduziert werden.

Durch eine zweckmäßige Kombination bekannter Korrelations-Methoden wurde ein neues Verfahren zur *merkmalsbasierten Korrespondenz-Analyse* vorgestellt. Nach einem Verträglichkeitstest mit differentiellen Invarianten und einer geometrischen Anpassung mit der Kleinste-Quadrate-Bildzuordnung kann die Ähnlichkeit von Merkmalen durch die normierte Kreuz-Korrelation definiert werden. Die Leistungsfähigkeit der entwickelten Methode in Verbindung mit einer hierarchischen Bildzuordnungs-Strategie wurde anhand mehrerer Beispiele demonstriert. Hier wäre ein direkter Vergleich mit dem Korrespondenz-Verfahren von MATAS u.a. [MAT02] interessant, das recht gute Ergebnisse durch die Zuordnung von *maximal stabilen Extremal-Regionen* (MSER) verspricht.

Für die automatische Erkennung und Beseitigung von falsch zugeordneten Merkmalen wurde ein neues Verfahren entwickelt, das auf der Grundlage eines genetischen Algorithmus die *robuste Parameterschätzung* signifikant beschleunigt. Durch die Auswertung der *Trifokal-Geometrie* wird die Anzahl der fehlerhaft zugeordneten Punktmerkmale noch weiter reduziert. Um aus den homologen Punkten schließlich eine projektive Rekonstruktion zu erhalten, wurde eine stabile Methode zur Bestimmung von räumlichen Koordinaten durch *Triangulation* beschrieben.

Des Weiteren wurden zwei Näherungsverfahren analysiert, die einzelne Stereobilder in einem gemeinsamen Bildverband kombinieren. Durch den Einsatz einer *projektiven Bündel-Ausgleichung* kann anschließend eine spannungsfreie Aufnahmegeometrie erreicht werden. Das bei der Implementierung verwendete Optimierungsverfahren sollte allerdings bei einer größeren Anzahl von Bildern oder Punkten durch einen effizienteren Algorithmus ersetzt werden [TRI99, BAR01].

Für die *Selbst-Kalibrierung* fokussierbarer Kameras wurde ein Verfahren auf der Basis des dualen absoluten Quadric untersucht und dessen Ergebnisse durch Eigenschaften der essentiellen Matrix verbessert. Neben der Behandlung von kritischen Oberflächen [POL02] und Aufnahmekonfigurationen [STU97, TRI97] sollte noch die robuste Eigen-system-Zerlegung mit dem komplexwertigen Ansatz von SEO und HONG [SEO01] integriert werden.

Um den Aufwand bei der Erstellung von Oberflächenmodellen zu reduzieren, wurde ein hybrider Ansatz zur *Normalbild-Erstellung* beschrieben, der durch die Verwendung von Polarkoordinaten beliebige Bildanordnungen im Nahbereich erlaubt. Als abschließendes Anwendungsbeispiel für eine *3D-Rekonstruktion* wurde erläutert, wie ein Objekt durch Ausschneiden seiner Konturen aus einem Volumen modelliert werden kann. Dabei konnte gezeigt werden, dass der erforderliche Berechnungsaufwand durch eine hierarchische Verfeinerung mit einer Octree-Struktur deutlich reduziert wird.

In der vorliegenden Arbeit konnte nachgewiesen werden, dass mit den beschriebenen Verfahren eine automatische Auswertung von Bildern im Nahbereich auch ohne Zusatzinformation möglich ist. Durch diese Entwicklung können die photogrammetrischen Methoden in vielen Anwendungsbereichen noch flexibler und leistungsfähiger eingesetzt werden.

## Anhang A

### A.1 Singulärwert-Zerlegung

Die *Singulärwert-Zerlegung* (SVD) ist eine Faktorisierung einer Matrix zu  $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} u_{11} & \cdots & u_{1n} \\ u_{21} & \cdots & u_{2n} \\ \vdots & \ddots & \vdots \\ u_{m1} & \cdots & u_{mn} \end{bmatrix} \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix}^T, \quad (\text{A.1})$$

wobei  $\mathbf{U}$  eine  $m \times n$ -Matrix mit orthogonalen Spalten,  $\mathbf{D}$  eine  $n \times n$ -Diagonal-Matrix mit positiven Elementen und  $\mathbf{V}$  eine Orthogonal-Matrix ist. Die Diagonalelemente von  $\mathbf{D}$  sind die so genannten *Singulärwerte* der Matrix  $\mathbf{A}$  und entsprechen den Quadratwurzeln der Eigenwerte von  $\mathbf{A}^T \mathbf{A}$ .

Für die praktische Implementierung wurde die SVD aus den ‚*Numerical Recipes in C++*‘ [PRE02] eingesetzt, wobei zunächst  $m \geq n$  vorausgesetzt wird. In den anderen Fällen kann durch Auffüllen von zusätzlichen Zeilen mit Nullen sehr einfach eine quadratische Matrix erzeugt werden. Des Weiteren wurde die Dekomposition so erweitert, dass die Diagonalelemente von  $\mathbf{D}$  in absteigender numerischer Reihenfolge sortiert werden  $d_1 \geq \dots \geq d_n$ . Damit entspricht die Position der letzten Spalte von  $\mathbf{V}$  immer dem kleinsten Singulärwert von  $\mathbf{A}$ . Eine Ausnahme bildet die Lösung inhomogener linearer Gleichungssysteme (vgl. Anhang A.2.2).

### A.2 Lineare Kleinste-Quadrate-Lösung

#### A.2.1 Homogene Gleichungssysteme

Bei linearen homogenen Gleichungen des Typs

$$\mathbf{A} \mathbf{x} = \mathbf{0} \quad (\text{A.2})$$

soll ein Gleichungssystem mit mindestens so vielen Gleichungen wie Unbekannten betrachtet werden und die nahe liegende Lösung  $\mathbf{x} = \mathbf{0}$  ausgeschlossen sein. Als geeignete Bedingung kann die Normierung mit  $\|\mathbf{x}\| = 1$  durchgeführt werden, sofern der Lösungsvektor  $\mathbf{x}$  nur bis auf einen Skalierungsfaktor bestimmt werden muss. Die *Singulärwert-Zerlegung* (vgl. Anhang A.1) ist eine verbreitete Methode bei der Lösung von überbestimmten Gleichungssystemen.

Eine gute Schätzung der gesuchten Kleinste-Quadrate-Lösung für  $\mathbf{x}$  ist der Rechtseigenvektor, der dem kleinsten Eigenwert von  $\mathbf{A}$  entspricht. Nach einer SVD mit  $\mathbf{A} = \mathbf{UDV}^T$  ist dieser durch die letzten Spalte von  $\mathbf{V}$  gegeben. Eine wichtige Voraussetzung für die stabile Lösung von homogenen Gleichungssystemen ist allerdings eine entsprechende Normierung der Daten [HAR97a].

## A.2.2 Inhomogene Gleichungssysteme

Eine übliche Methode zur Berechnung von überbestimmten linearen Gleichungssystemen der Form

$$\mathbf{A} \mathbf{x} = \mathbf{b} \quad (\text{A.3})$$

besteht in der Verwendung von Normalgleichungen. Wenn beide Seiten der Gleichung mit der Transponierten von  $\mathbf{A}$  multipliziert werden

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b} \quad (\text{A.4})$$

und  $\mathbf{A}$  den vollen Rang besitzt, dann ist die quadratische Matrix  $\mathbf{A}^T \mathbf{A}$  invertierbar. Die gesuchte Lösung für  $\mathbf{x}$ , bei der  $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|$  minimal ist, kann schließlich mit Hilfe der MOORE-PENROSE Pseudo-Inversen

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (\text{A.5})$$

gefunden werden. Die Kleinste-Quadrate-Lösung für  $\mathbf{x}$  lässt sich aber auch mit Hilfe der Singulärwert-Zerlegung bestimmen. HARTLEY [HAR00] schlägt zur Berechnung von  $\mathbf{x}$  den folgenden Algorithmus vor:

1. Bestimmen der SVD  $\mathbf{A} = \mathbf{UDV}^T$
2. Definieren von  $\mathbf{b}' = \mathbf{U}^T \mathbf{b}$
3. Bestimmen des Vektors  $\mathbf{y}$  durch  $y_i = b'_i / d_i$ ,  
wobei  $d_i$  das  $i$ -te Diagonalelement von  $\mathbf{D}$  ist. (A.6)
4. Die Lösung ist  $\mathbf{x} = \mathbf{V}\mathbf{y}$ .

Die normalerweise für die Singulärwert-Zerlegung zweckmäßige Sortierung der Diagonalelemente von  $\mathbf{D}$  darf bei diesem Verfahren allerdings nicht durchgeführt werden (vgl. Anhang A.2.1). Diese etwas aufwändigere Lösung mit Hilfe der SVD ist der Pseudo-Inversen vorzuziehen, da sie numerisch stabiler ist und der Algorithmus theoretisch nicht fehlschlagen kann [PRE02].

## A.3 Robuste Schätzverfahren

In der Regel wird angenommen, dass die gemessenen Daten (z.B. die homologen Bildpunkte oder 3D-Passpunkte) nur kleine Abweichungen enthalten und somit einer GAUSSschen Fehlerverteilung entsprechen. Bei einer automatisierten Auswertung können die Messwerte allerdings auch durch grobe Fehler verfälscht werden.

Ein besonderes Problem sind hier die so genannten *Hebelpunkte*, die weit entfernt von der Masse der übrigen Daten liegen und einen erheblichen Einfluss auf die Parameterschätzung ausüben [KOC97]. Durch den Einsatz einer robusten Statistik können solche Ausreißer identifiziert und ihr negativer Einfluss auf die Lösung ausgeschlossen werden.

### A.3.1 M-Schätzer (HUBER, TUKEY)

Um eine Parameterschätzung unempfindlich gegenüber Ausreißern durchzuführen, werden häufig robuste *Maximum-Likelihood-Schätzer* (M-Schätzer) verwendet. Sei  $e_i$  der Fehler der  $i$ -ten Beobachtung, dann minimiert die Methode der kleinsten Quadrate üblicherweise die Summe der quadratischen Residuen  $e_i^2$ . Das Ziel der M-Schätzer besteht nun darin, dieses empfindliche Maß durch eine geeignete Einflussfunktion zu ersetzen

$$C = \sum_i \rho(e_i), \quad (\text{A.7})$$

wofür in der Literatur viele Lösungsvorschläge existieren. Eine umfassende Übersicht der verschiedenen Methoden ist z.B. in ZHANG [ZHA98] gegeben. Für die Implementierung wurden die HUBER- und TUKEY-Funktionen ausgewählt.

#### A.3.1.1 Minmax-Funktion von HUBER

Eine sehr populäre und einfache Variante ist die robuste Minmax-Funktion von HUBER

$$\rho(e) = \min(t, \max(e, -t)), \quad (\text{A.8})$$

welche die Residuen  $e$  mit sehr großem Fehler von der Schätzung ausschließt, indem der Einfluss ab einem bestimmten Schwellwert  $t$  auf ein konstantes Maß beschränkt wird (vgl. Abb. A.1).

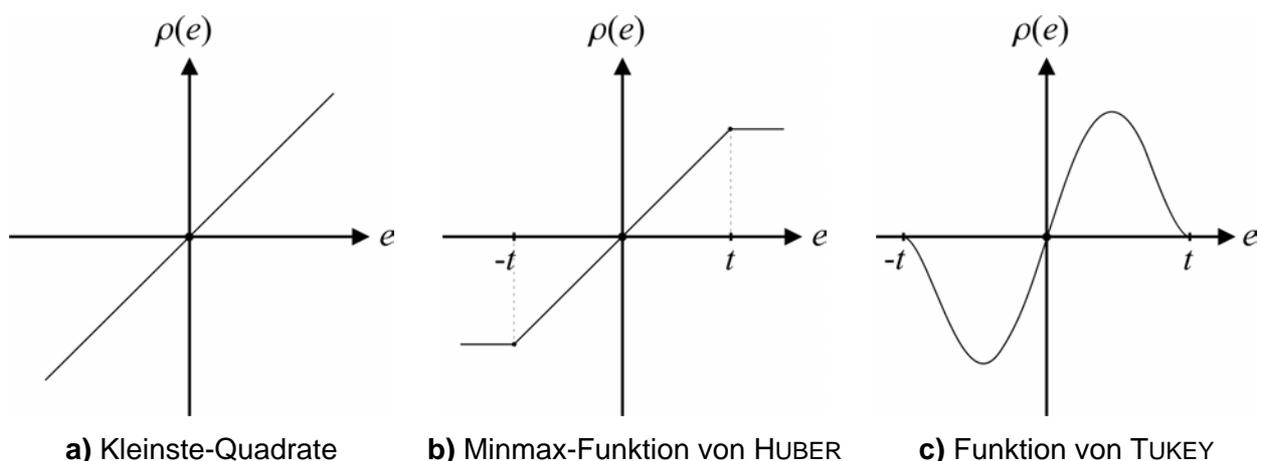


Abb. A.1: Einflussfunktionen für Maximum-Likelihood-Schätzer

Dimension	Modell	Schwellwert $t$
1	Fundamental-Matrix	$1,96\sigma$
2	2D-Homographie	$2,45\sigma$
3	3D-Homographie, Trifokal-Geometrie	$2,79\sigma$

Tab. A.1: Schwellwerte bei einer Zuverlässigkeit von 95 Prozent

### A.3.1.2 TUKEY Funktion

Bei der Funktion von TUKEY ist der Einfluss nicht nur nach oben beschränkt, sondern er reduziert sich sogar ab einem bestimmten Wert wieder

$$\rho(e) = \begin{cases} e(t^2 - e^2)^2, & |e| < t \\ 0, & |e| \geq t \end{cases} \quad (\text{A.9})$$

Der Einfluss der beschriebenen Funktionen ist in Abb. A.1 veranschaulicht.

Der Schwellwert  $t$  kann entweder empirisch bestimmt werden oder mit Hilfe der  $\chi^2$ -Verteilung geschätzt werden. Unter der gebräuchlichen Annahme, dass die Wahrscheinlichkeit für eine korrekte Datenauswahl 95 Prozent beträgt, sind einige Schwellwerte in Tab. A.1 aufgeführt.

Um den Einfluss von Ausreißern auch bei der Berechnung von der Standardabweichung  $\sigma$  zu reduzieren, schlägt ROUSSEEUW [ROU87] eine robuste Methode vor. Anstelle der durchschnittlichen Abweichung der Residuen  $e$  vom Mittelwert wurde bei der Implementierung die folgende *robuste Standardabweichung* eingesetzt:

$$\sigma = 1,4826 \cdot \left(1 + \frac{5}{n-p}\right) \cdot \text{median}_i |e_i|, \quad (\text{A.10})$$

wobei  $n$  die Anzahl der Beobachtungen und  $p$  die Größe des Parameterraums bezeichnen. Der Median bezeichnet dabei den Fehler des mittleren Elements aus der sortierten Liste der Residuen.

### A.3.2 Verfahren des Least-Median-of-Squares (LMedS)

Eine weitere robuste Methode zur Parameterschätzung stellt das Verfahren des *Least-Median-of-Squares* (LMedS) dar, dessen Fehlerfunktion folgendermaßen definiert ist:

$$C = \text{median}_i e_i^2 \quad (\text{A.11})$$

Die Verwendung des äußerst stabilen Medians anstelle einer Summe toleriert bis zu 50 Prozent Ausreißer, ohne dass die Schätzung zusammenbricht und es braucht außerdem kein Schwellwert definiert zu werden.

### A.3.3 Monte-Carlo-Methode (RANSAC)

Im Prinzip verwenden die klassisch glättend wirkenden Methoden möglichst viele Daten für eine erste Schätzung, um danach Ausreißer zu erkennen und zu eliminieren. Im Gegensatz dazu verwendet die Methode von FISCHLER und BOLLES [FIS81] möglichst wenige Daten für eine erste Schätzung und maximiert anschließend die Menge der zu diesem Modell konsistenten Daten.

Für die robuste Bestimmung der Fundamental-Matrix  $\mathbf{F}$  hat sich dieses RANSAC-Verfahren (*RANdom SAMpling Consensus*) bewährt [TOR95]. Diese Vorgehensweise ist jedoch sehr rechenintensiv, denn für  $n$  Beobachtungen müsste bei  $k$  unbekanntem Parametern eine Kombination von  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$  Möglichkeiten berechnet werden. Zur Einschränkung des Rechenaufwandes wird daher nur eine Teilmenge zufällig ausgewählt und mit Hilfe der Wahrscheinlichkeitsrechnung die Anzahl der notwendigen Parameterschätzungen reduziert.

#### A.3.3.1 Optimierung durch zufälliges Probieren

Der Algorithmus generiert zunächst ein Modell, indem aus der vorhandenen Datenmenge  $S$  mit  $n$  Punkten ein minimaler Datensatz mit  $s$  Punkten zufällig ausgewählt und dafür eine Lösung bestimmt wird (z.B. mit dem *7-Punkt-Algorithmus* Abschnitt 4.3.4). Dabei sollten degenerierte Konfigurationen (z.B. kollineare Punkte) erkannt und zur Vermeidung von Extrapolationsproblemen auf eine gute räumliche Verteilung geachtet werden.

Anschließend kann mit einer robusten Fehlerfunktion (z.B. dem *M-Schätzer* Gleichung A.7 oder dem *LMedS-Verfahren* Gleichung A.11) der Fehler des geschätzten Modells bezüglich aller Punkte in  $S$  bestimmt und iterativ minimiert werden. Die zufällige Auswahl eines minimalen Datensatzes wird dabei solange wiederholt, bis ein geeignetes Abbruchkriterium erfüllt ist. Bei dieser Art der Kosten-Minimierung werden allerdings viele Daten als Ausreißer deklariert.

Eine alternative Lösung besteht darin, die Anzahl der zum Modell konsistenten Daten

$$\rho(e) = \begin{cases} 1, & |e| < t \\ 0, & |e| \geq t \end{cases} \quad (\text{A.12})$$

zu maximieren. Sollten dabei mehrere Lösungen in Frage kommen, so wird der Datensatz mit der kleinsten Standardabweichung  $\sigma$  ausgewählt. Nach der Optimierung werden die Ausreißer eliminiert und das Modell mit allen konsistenten Daten neu geschätzt.

#### A.3.3.2 Adaptives Abbruchkriterium

Unter Berücksichtigung des Berechnungsaufwandes ist es nicht sinnvoll, alle möglichen Datenkombinationen zu testen. Bei der Maximierung der konsistenten Daten lässt sich aber die Anzahl der notwendigen Iterationen folgendermaßen abschätzen. Sei

$$\varepsilon = 1 - \frac{C}{n} \quad (\text{A.13})$$

der Anteil der Ausreißer in der Datenmenge  $S$  und

$$p = 1 - \left(1 - (1 - \varepsilon)^s\right)^N \quad (\text{A.14})$$

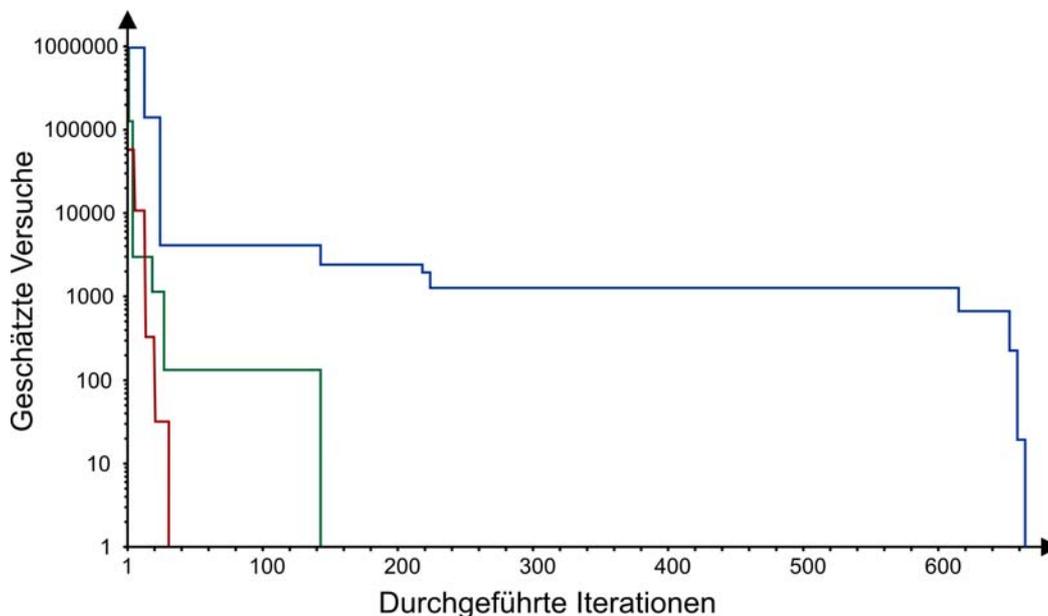
die Sicherheit, dass sich mindestens eine minimale Auswahl mit  $s$  Elementen ohne Ausreißer unter den  $N$  Datensätzen befindet, dann lässt sich die Mindestanzahl der Versuche folgendermaßen abschätzen:

$$N = \frac{\ln(1-p)}{\ln\left(1 - (1-\varepsilon)^s\right)} \quad (\text{A.15})$$

Als Beispiel reichen bei einer Tolerierung von 45 Prozent Ausreißern ( $\varepsilon = 0,45$ ) und einer Sicherheit von 99 Prozent ( $p = 0,99$ ) für die lineare Bestimmung der Fundamental-Matrix ( $s = 8$ ) bereits  $N = 548$  statt  $\binom{25}{8} = 1.081.575$  Versuche aus.

Bei dem *adaptiven Verfahren* wird die Anzahl der notwendigen Iterationen  $N$  bei jedem Durchlauf mit dem aktuellen  $\varepsilon$  neu bestimmt, so dass  $N$  sich in großen Schritten reduziert. Abb. A.2 illustriert die dynamische Anpassung von  $N$  an drei typischen Beispielen.

Durch die Verwendung der Wahrscheinlichkeitsrechnung kann es aber in einigen ungünstigen Fällen dazu kommen, dass die Anzahl der konsistenten Daten zu gering ist. Daher ist der Einsatz eines zusätzlichen Abbruchkriteriums zu empfehlen, das die Anzahl der Ausreißer auf maximal 50 Prozent beschränkt. Die praktischen Erfahrungen mit stark gestörten Daten haben auch gezeigt, dass die geschätzte Mindestanzahl häufig viel zu gering ist und wesentlich mehr Versuche notwendig sind.



**Abb. A.2:** Adaptive Anpassung des Abbruchkriteriums

## Literaturverzeichnis

- [ABD71] Y.I. ABDEL-AZIZ und H.M. KARARA: *Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry*, Proc. ASP Symposium on Close-Range Photogrammetry, University of Illinois, Urbana-Champaign, S. 1-18, Jan. 1971.
- [ACK84] F. ACKERMANN: *High precision digital image correlation*, Schriftenreihe des Instituts für Photogrammetrie, Universität Stuttgart, Heft 9, S. 231-243, 1984.
- [ALB01] J. ALBERTZ: *Einführung in die Fernerkundung – Grundlagen der Interpretation von Luft- und Satellitenbildern*, 2. Auflage, Wissenschaftliche Buchgesellschaft, ISBN: 3-534-14624-7, 250 S., 2001.
- [ALB97] J. ALBERTZ und A. WIEDEMANN (Hrsg.): *Architekturphotogrammetrie gestern-heute-morgen*, MEYDENBAUER-Kolloquium, Technische Universität Berlin, ISBN: 3-7983-1700-3, 130 S., 1997.
- [BAR01] A. BARTOLI und P. STURM: *Three new algorithms for projective bundle adjustment with minimum parameters*, TR 4236, INRIA, Sophia Antipolis, Frankreich, 28 S., 2001.
- [BRA96] G. BRANDSTÄTTER: *Fundamentals of algebro-projective photogrammetry*, In: Sitzungsberichte, Math.-naturwiss. Klasse Abt. II, Math., Physik. und Techn. Wissenschaften, Österreichische Akademie der Wissenschaften, Bd. II, Nr. 205, S. 57-109, 1996.
- [BRA91] R. BRAUSE: *Neuronale Netze – eine Einführung in die Neuroinformatik*, Teubner, Stuttgart, ISBN: 3-519-02247-8, 291 S., 1991.
- [CAN86] J. CANNY: *A computational approach to edge detection*, IEEE T-PAMI, Bd. 8, Nr. 6, S. 679-698, Nov. 1986.
- [CAR98] S. CARLSSON und D. WEINSHALL: *Dual computation of projective shape and camera positions from multiple images*, IJCV, Bd. 27, Nr. 3, S. 227-241, 1998.
- [CAS96] K. CASTLEMAN: *Digital image processing*, Prentice-Hall, New Jersey, USA, ISBN: 0-13-398058-8, 667 S., 1996.

- [CRI99] A. CRIMINISI: *Accurate visual metrology from single and multiple uncalibrated images*, Dissertation, Robotics Research Group, Universität Oxford, 210 S., 1999.
- [CSU99] G. CSURKA, D. DEMIRDJIAN und R. HORAUD: *Finding the collineation between two projective reconstructions*, CVIU, Bd. 75, Nr. 3, S. 260-268, 1999.
- [CSU97] G. CSURKA, C. ZELLER, Z. ZHANG und O. FAUGERAS: *Characterizing the uncertainty of the fundamental matrix*, CVGIP - Image Understanding, Bd. 68, Nr. 1, S. 18-36, 1997.
- [DAS49] G.B. DAS: *A mathematical approach to problems in photogrammetry*, Empire Survey Review, Bd. X, Nr. 73, Juli 1949.
- [DER93] R. DERICHE und G. GIRAUDON: *A computational approach for corner and vertex detection*, IJCV, Bd. 10, Nr. 2, S. 101-124, 1993.
- [DER90] R. DERICHE: *Fast algorithms for low-level vision*, IEEE T-PAMI, Bd. 12, S. 78-87, 1990.
- [FAU01] O.D. FAUGERAS und Q.-T. LUONG: *The geometry of multiple images*, The MIT Press, Cambridge, Massachusetts, ISBN: 0-262-06220-8, 644 S., 2001.
- [FAU96] O.D. FAUGERAS und L. ROBERT: *What can two images tell us about a third one?*, IJCV, Bd. 18, Nr. 1, S. 5-20, Apr. 1996.
- [FAU92] O.D. FAUGERAS: *What can be seen in three dimensions with an uncalibrated stereo rig?*, In G. SANDINI (Hrsg.): Proc. 2<sup>nd</sup> ECCV'92, LNCS 588, Springer-Verlag, Santa Margherita Ligure, Italien, S. 563-578, Mai 1992.
- [FIN94] G.D. FINLAYSON, M.S. DREW und B. FUNT: *Color constancy – Generalized diagonal transforms suffice*, Journal of the Optical Society of America, Bd. A, Nr. 11, S. 3011-3019, Nov. 1994.
- [FIS81] M.A. FISCHLER und R.C. BOLLES: *Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography*, CVGIP 17, S. 173-180, 1981.
- [FIT98] A.W. FITZGIBBON, G. CROSS und A. ZISSERMANN: *Automatic 3D model construction for turntable sequences*, In: R. KOCH und L. VAN GOOL (Hrsg.): *3D structure from multiple images of large-scale environments SMILE'98*, LNCS 1506, Springer-Verlag, S. 155-170, Juni 1998.
- [FÖR00] W. FÖRSTNER: *Moderne Orientierungsverfahren*, Photogrammetrie – Fernerkundung – Geoinformation PFG, Heft 3, Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, S. 163-176, 2000.

- [FÖR94] W. FÖRSTNER: *A framework for low level feature extraction*, In: J.O. ECKLUNDH (Hrsg.): Proc. 3<sup>rd</sup> ECCV'94, LNCS 800, Springer-Verlag, Stockholm, Schweden, S. 383-394, Mai 1994.
- [FÖR93] W. FÖRSTNER: *Image matching*, In: R.M. HARALICK und L.G. SHAPIRO (Hrsg.) *Computer and robot vision*, Bd. 2, Addison-Wesley, Reading, Massachusetts, USA, ISBN: 0-201-56943-4, S. 289-362, 1993.
- [FÖR87] W. FÖRSTNER und E. GÜLCH: *A fast operator for detection and precise location of distinct points, corners and circular features*, Proc. Intercommission Conf. on Fast Processing of Photogrammetric Data, Interlaken, Schweiz, S. 281-305, 1987.
- [FÖR85] W. FÖRSTNER: *Prinzip und Leistungsfähigkeit der Korrelation und Zuordnung digitaler Bilder*, Schriftenreihe Institut für Photogrammetrie, Universität Stuttgart, Heft 11, Stuttgart, S. 69-90, 1985.
- [FUS00] A. FUSIELLO: *Uncalibrated euclidean reconstruction – A review*, Image and Vision Computing 18, Elsevier Verlag, S. 555-563, 2000.
- [GON01] R.C. GONZALEZ und R.E. WOODS: *Digital image processing*, 2. Auflage, Prentice-Hall, New Jersey, USA, ISBN: 0-201-18075-8, 793 S., 2001.
- [HAR00] R. HARTLEY und A. ZISSERMAN: *Multiple view geometry in computer vision*, Cambridge University Press, ISBN: 0-521-62304-9, 607 S., 2000.
- [HAR99] R. HARTLEY: *Theory and practice of projective rectification*, IJCV, Bd. 35, Nr. 2, S. 1-16, 1999.
- [HAR98] R. HARTLEY: *Chirality*, IJCV, Bd. 26, Nr. 1, S. 41-61, 1998.
- [HAR97a] R. HARTLEY: *In defense of the eight-point algorithm*, IEEE T-PAMI, Bd. 19, Nr. 6, S. 580-593, Okt. 1997.
- [HAR97b] R. HARTLEY: *Kruppa's equations derived from the fundamental matrix*, IEEE T-PAMI, Bd. 19, S. 133-135, Feb. 1997.
- [HAR95] R. HARTLEY und P. STURM: *Triangulation*, Proc. 6<sup>th</sup> Int. Conf. on CAIP'95, Prag, Tschechische Republik, S. 190-197, 1995.
- [HAR92a] R.M. HARALICK und L.G. SHAPIRO: *Computer and Robot Vision*, Addison-Wesley, Reading, Massachusetts, USA, Bd.1: ISBN: 0-201-10877-1, 1992 und Bd. 2: ISBN: 0-201-56943-4, 630 S., 1993.
- [HAR92b] R. HARTLEY: *Estimation of relative camera positions for uncalibrated cameras*, In G. SANDINI (Hrsg.): Proc. 2<sup>nd</sup> ECCV'92, LNCS 588, Springer-Verlag, Santa Margherita Ligure, Italien, S. 579-587, Mai 1992.
- [HAR88] C. HARRIS und M. STEPHENS: *A combined corner and edge detector*, Proc. 4<sup>th</sup> Alvey Vision Conf., S. 147-151, Aug. 1988.

- [HEI95] C. HEIPKE: *Digitale photogrammetrische Arbeitsstationen*, Habilitation, DGK, Reihe C, Nr. 450, Beck'sche Verlagsbuchhandlung, München, ISBN: 3-7696-9493-7, 111 S., 1995.
- [HEI90] C. HEIPKE: *Integration von Bildzuordnung, Punktbestimmung, Oberflächenrekonstruktion und Orthophotoprojektion innerhalb der digitalen Photogrammetrie*, Dissertation, DGK, Reihe C, Nr. 366, Beck'sche Verlagsbuchhandlung, München, 89 S., 1990.
- [HEY97] A. HEYDEN und K. ÅSTRÖM: *Euclidean reconstruction from image sequences with varying and unknown focal length and principal point*, Proc. Int. Conf. on CVPR'97, IEEE Computer Society Press, San Juan, Puerto Rico, S. 438-443, 1997.
- [HIL90] D. HILBERT: *Theory of algebraic invariants*, Cambridge Mathematica Library, Cambridge University Press, 1890.
- [HOL75] J.H. HOLLAND: *Adaption in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*, University of Michigan Press, Ann Arbor, ISBN 0-472-08460-7, 206 S., 1975.
- [JÄH97] B. JÄHNE: *Practical handbook on image processing*, CRC-Press, Boca Raton, Florida, USA, ISBN: 0-8493-8906-2, 589 S., 1997.
- [KLE96a] R. KLETTE und P. ZAMPERONI: *Handbook of image processing operators*, J. Wiley & Sons, New York, ISBN: 0471-95642-2, 416 S., 1996.
- [KLE96b] R. KLETTE, A. KOSCHAN und K. SCHLÜNS: *Computer Vision – Räumliche Information aus digitalen Bildern*, Vieweg Verlag, Braunschweig, ISBN: 3-528-06625-3, 382 S., 1996.
- [KOC97] K.R. KOCH: *Parameterschätzung und Hypothesentests in linearen Modellen*, 3. Auflage, Dümmlers Verlag, ISBN: 3-427-78923-3, 384 S., 1997.
- [KÖT03] U. KÖTHE: *Edge and junction detection with an improved structure tensor*, In B. MICHAELIS, G. KRELL (Hrsg.): Proc. of 25<sup>th</sup> DAGM Symposium, Magdeburg, LNCS 2781, Springer-Verlag, S. 25-32, 2003.
- [KOS98] A. KOSCHAN: *Digitale Farbbildverarbeitung – Erfordernisse und Möglichkeiten*, Habilitation, Fachbereich Informatik, Technische Universität Berlin, 1998.
- [KOS96] A. KOSCHAN, V. RODEHORST und K. SPILLER: *Color stereo vision using hierarchical block matching and active color illumination*, Proc. 13<sup>th</sup> ICPR'96, IEEE Computer Society Press, Wien, Österreich, Bd. 1, S. 835-839, Aug. 1996.

- [KOS95] A. KOSCHAN und V. RODEHORST: *Towards real-time stereo employing parallel algorithms for edge-based and dense stereo matching*, Proc. Workshop on CAMP'95, IEEE Computer Society Press, Como, Italien, S. 234-241, Sep. 1995.
- [KOS91] A. KOSCHAN: *Eine Methodenbank zur Evaluierung von Stereo-Vision-Verfahren*, Dissertation, TR 91-09, Fachbereich Informatik, Technische Universität Berlin, 242 S., Juni 1991.
- [KRU13] E. KRUPPA: *Zur Ermittlung eines Objektes aus zwei Perspektiven mit innerer Orientierung*, Sitzungsberichte Österreichische Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse, Abteilung IIa, 122, S. 1939-1948, 1913.
- [KUZ02] Y. KUZU: *Photorealistic object reconstruction using color image matching*, Proc. of ISPRS Symp. Commission V – Close Range Visualization Techniques, Bd. XXXIV, Nr. 5, Korfu, Griechenland, S. 169-174, Sep. 2002.
- [KUZ01] Y. KUZU und V. RODEHORST: *Volumetric modeling using shape from silhouette*, In M.O. ALTAN und L. GRÜNDIG (Hrsg.): Proc. 4<sup>th</sup> Turkish-German Joint Geodetic Days TGJGD'01, Bd. 1, Berlin, S. 469-476, Apr. 2001.
- [LON81] H.C. LONGUET-HIGGINS: *A computer algorithm for reconstructing a scene from two projections*, Nature 293, S. 133-135, Sep. 1981.
- [LOU99] M. LOURAKIS und R. DERICHE: *Camera self-calibration using the singular value decomposition of the fundamental matrix: from point correspondences to 3D measurements*, TR 3748, INRIA, Sophia Antipolis, Frankreich, 45 S., Aug. 1999.
- [LUH02] T. LUHMANN: *Nahbereichsphotogrammetrie in der Praxis – Beispiele und Lösungen*, Wichmann Verlag, Heidelberg, ISBN: 3-87907-385-6, 318 S., 2002.
- [LUH00] T. LUHMANN: *Nahbereichsphotogrammetrie – Grundlagen, Methoden und Anwendungen*, Wichmann Verlag, Heidelberg, ISBN: 3-87907-321-X, 571 S., 2000.
- [LUO97] Q.-T. LUONG und O.D. FAUGERAS: *Self-calibration of a moving camera from point correspondences and fundamental matrices*, IJCV, Bd. 22, Nr. 3, S. 261-289, 1997.
- [LUO96] Q.-T. LUONG und T. VIÉVILLE: *Canonical representations for the geometries of multiple projective views*, CVIU, Bd. 64, Nr. 2, S. 193-229, 1996.
- [MAL98] H.A. MALLOT: *Sehen und die Verarbeitung visueller Informationen*, Verlag Vieweg, Braunschweig, ISBN: 3-528-05659-2, 258 S., 1998.

- [MAN92] R. MANMATHA und J. OLIENSIS: *Measuring the affine transformation: Recovering scale and rotation*, COINS TR92-74, Computer Science Department, University of Massachusetts, Amherst, USA, 33 S., 1992.
- [MAR80] D. MARR und E. HILDRETH: *Theory of edge detection*, Proc. Royal Society of London, Bd. 207, S.187-217, 1980.
- [MAR83] W.N. MARTIN und J.K. AGGARWAL: *Volumetric descriptions of objects from multiple views*, IEEE T-PAMI, Bd. 5, Nr. 2, S. 150-158, 1983.
- [MAT02] J. MATAS, O. CHUM, M. URBAN und T. PAJDLA: *Robust wide baseline stereo from maximally stable extremal regions*, In D. MARSHALL und P.L. ROSIN (Hrsg.): Proc. 13<sup>th</sup> British Machine Vision Conf. BMVC'02, Cardiff, UK, S. 384-393, Sep. 2002.
- [MAY02] H. MAYER, W. BÜCHNER, T. RIEGEL: *View interpolation with the trifocal tensor from uncalibrated imagery*, Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Bd. 34, Nr. 5, S. 192-197, 2002.
- [MAY93] S.J. MAYBANK: *Theory of reconstruction from image motion*, Series in Information Sciences, Bd. 28, Springer-Verlag, ISBN 3-5405-5537-4, 261 S., 1993.
- [MAY92] S.J. MAYBANK und O.D. FAUGERAS: *A theory of self-calibration of a moving camera*, IJCV, Bd. 8, Nr. 2, S. 123-151, 1992.
- [MEH94] D. MEHREN und V. RODEHORST: *Gestaltsanalyse komplexer Objekte bei kontrollierter Bewegung*, CV-Bericht 27, Diplomarbeit, Fachbereich Informatik, Technische Universität Berlin, 108 S., 1994.
- [MEN99] P. MENDONCA und R. DERICHE: *A simple technique for self-calibration*, Proc. Int. Conf. on CVPR'99, Fort Collins, Colorado, USA, ISBN 0-7695-0149-4, S. 112-116, Juni 1999.
- [MIN01] F. MINDRU, T. TUYTELAARS, T. MOONS und L. VAN GOOL: *Invariant color features for matching and target recognition*, In BALTSAVIAS u.a. (Hrsg.): Automatic Extraction of Man-Made Objects from Aerial and Space Images, Bd. III, Swets & Zeitlinger, Lisse, ISBN: 90-5809-252-6, S. 35-45, 2001.
- [MON98] P. MONTESINOS, V. GOUET und R. DERICHE: *Differential invariants for color images*, Proc. 14<sup>th</sup> ICPR'98, IEEE Computer Society Press, Brisbane, Australien, 15 S., 1998.
- [MOR77] H. P. MORAVEC: *Toward automatic visual obstacle avoidance*, Proc. 5<sup>th</sup> Int. Joint Conf. Artificial Intelligence IJCAI'77, Cambridge, USA, S. 584, 1977.

- [ORA01] D.T. ORAM: *Rectification for any epipolar geometry*, Proc. British Machine Vision Conf. BMVC'01, London, UK, S. 653-662, Sep. 2001.
- [POL02] M. POLLEFEYS, F. VERBIEST und L. VAN GOOL: *Surviving dominant planes in uncalibrated structure and motion recovery*, In A. HEYDEN u.a. (Hrsg.): Proc. 7<sup>th</sup> ECCV'02, LNCS 2351, Springer Verlag, Kopenhagen, Dänemark, S. 837-851, Mai 2002.
- [POL99a] M. POLLEFEYS, R. KOCH und L. VAN GOOL: *A simple and efficient rectification method for general motion*, Proc. 7<sup>th</sup> ICCV, IEEE Computer Society Press, Korfu, Griechenland, S. 496-501, Sep. 1999.
- [POL99b] M. POLLEFEYS: *Self-calibration and metric 3D reconstruction from uncalibrated image sequences*, Dissertation, Katholische Universität Leuven, 227 S., Mai 1999.
- [POL99c] M. POLLEFEYS, R. KOCH und L. VAN GOOL: *Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters*, IJCV, Bd. 32, Nr. 1, S. 7-25, 1999.
- [PON98] J. PONCE und Y. GENC: *Epipolar geometry and linear subspace methods: A new approach to weak calibration*, IJCV, Bd. 28, Nr. 3, S. 223-243, 1998.
- [PRE02] W.H. PRESS, B.P. FLANNERY, S.A. TEUKOLSKY und W.T. VETTERLING: *Numerical recipes in C++*, 2. Ausgabe, Cambridge University Press, ISBN: 0-521-75033-4, 1002 S., 2002.
- [QUA94] L. QUAN: *Invariants of 6 points from 3 uncalibrated images*, In: J.O. ECKLUNDH (Hrsg.): Proc. 3<sup>rd</sup> ECCV'94, LNCS 800, Springer-Verlag, Stockholm, Schweden, S. 459-469, Mai 1994.
- [RAD93] B. RADIG: *Verarbeiten und Verstehen von Bildern*, In A. ENDRES (Hrsg.): *Handbuch der Informatik*, Oldenbourg Verlag, München, ISBN: 3-486-20787-3, 294 S., 1993.
- [REC73] I. RECHENBERG: *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Frommann-Holzboog Verlag, Stuttgart, ISBN 3-7728-0-374, 170 S., 1973.
- [ROS92] L. ROSENTHALER, F. HEITGER, O. KÜBLER und R. v.D. HEYDT: *Detection of general edges and keypoints*, In G. SANDINI (Hrsg.): Proc. 2<sup>nd</sup> ECCV'92, LNCS 588, Springer-Verlag, Santa Margherita Ligure, Italien, S. 78-86, Mai 1992.
- [ROU87] P.J. ROUSSEEUW und A.M. LEORY: *Robust regression and outlier detection*, Series in Probability and Statistics, J. Wiley & Sons, New York, 1987.

- [ROY97] S. ROY, J. MEUNIER und I. COX: *Cylindrical rectification to minimize epipolar distortion*, Proc. Int. Conf. on CVPR'97, IEEE Computer Society Press, San Juan, Puerto Rico, S. 393-399, Juni 1997.
- [RUS02] J.C. RUSS: *The image processing handbook*, 4. Auflage, CRC-Press, Boca Raton, Florida, USA, ISBN: 0-8493-1142-X, 732 S., 2002.
- [SAN08] H. V. SANDEN: *Die Bestimmung der Kernpunkte in der Photogrammetrie*, Dissertation, Universität Göttingen, Dez. 1908.
- [SCH03] T. SCHUMACHER: *Ausgleichungsansätze in der projektiven Geometrie*, Diplomarbeit, Fachgebiet Photogrammetrie und Kartographie, Technische Universität Berlin, 56 S., 2003.
- [SCH01] F. SCHAFFALITZKY und A. ZISSERMAN: *Viewpoint invariant texture matching and wide baseline stereo*, Proc. 8<sup>th</sup> ICCV, Bd. 2, IEEE Computer Society Press, Vancouver, Kanada, S. 636-643, Juli 2001.
- [SCH00] C. SCHMID, R. MOHR und C. BAUCKHAGE: *Evaluation of interest point detectors*, IJCV, Bd. 37, Nr. 2, S. 151-172, 2000.
- [SAM82] P. SAMPSON: *Fitting conic sections to 'very scattered' data – An iterative refinement of the Bookstein algorithm*, CVGIP, Bd. 18, S. 97-108, 1982.
- [SEM52] J.G. SEMPLE und G.T. KNEEBONE: *Algebraic projective geometry*, Oxford University Press, ISBN: 0-198-50363-6, 416 S., 1952 (Nachdruck 1979).
- [SEO01] Y. SEO und K.S. HONG: *A linear metric reconstruction by complex eigen-decomposition*, IEICE Transactions on Information and Systems, Bd. E84-D, Nr. 12, S. 1626-1632, Korea, Dez. 2001.
- [SMI97] S.M. SMITH und J. M. BRADY: *SUSAN – a new approach to low level image processing*, IJCV, Bd. 23, Nr. 1, S. 45-78, 1997.
- [SPE91] M.E. SPETSAKIS und Y. ALOIMONOS: *A multi-frame approach to visual motion perception*, IJCV, Bd. 6, S. 245-255, 1991.
- [SRI90] S.K. SRIVASTAVA und N. AHUJA: *Octree generation from object silhouettes in perspective views*, CVGIP, Nr. 49, S. 68-84, 1990.
- [STE98] C. STEGER: *Unbiased extraction of curvilinear structures from 2D and 3D images*, Dissertation, Fakultät für Informatik, Technische Universität München, Herbert Utz Verlag, München, ISBN 3-89675-346-0, 192 S., 1998.
- [STO91] J. STOLFI: *Orientated projective geometry*, Academic Press, Boston, USA, ISBN: 0-12-672025-8, 237 S., 1991.
- [STU97] P. STURM: *Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction*, Proc. Int. Conf. on CVPR'97, IEEE Computer Society Press, San Juan, Puerto Rico, S. 1100-1105, Juni 1997.

- [STU96] P. STURM und B. TRIGGS: *A factorized based algorithm for multi-image projective structure and motion*, In B. BUXTON und R. CIPOLLA (Hrsg.): Proc. 4<sup>th</sup> ECCV'96, LNCS 1065, Springer-Verlag, Cambridge, UK, S. 709-720, 1996.
- [SZE94] R. SZELISKI: *Image mosaicing for tele-reality applications*, Proc. IEEE Workshop on Applications of Computer Vision WACV'94, S. 230-236, 1994.
- [TAR90] H.D. TAGARE und R.J.P. DE FIGUEIREDO: *On the localization performance measure and optimal edge detection*, IEEE T-PAMI, Bd. 12, Nr. 12, S. 1186-1190, 1990.
- [TOR97] P.H.S. TORR und A. ZISSERMAN: *Robust parameterization and computation of the trifocal tensor*, Image and Vision Computing, Bd. 15, Nr. 8, S. 591-605, Aug. 1997.
- [TOR95] P.H.S. TORR: *Motion segmentation and outlier detection*, Dissertation, Dept. of Engineering Science, University of Oxford, 1995.
- [TRI99] B. TRIGGS, P. MCLAUCHLAN, R. HARTLEY und A. FITZGIBBON: *Bundle adjustment – A modern synthesis*, In B. TRIGGS u.a. (Hrsg.): Proc. Int. Workshop on Vision Algorithms – Theory and Practice, LNCS 1883, Springer-Verlag, Korfu, Griechenland, S. 298-372, Sep. 1999.
- [TRI97] B. TRIGGS: *Autocalibration and the absolute quadric*, Proc. Int. Conf. on CVPR'97, IEEE Computer Society Press, San Juan, Puerto Rico, S. 609-614, Juni 1997.
- [TRI95] B. TRIGGS: *The geometry of projective reconstruction: Matching constraints and the joint image*, Proc. 5<sup>th</sup> ICCV, IEEE Computer Society Press, Boston, USA, S. 338-343, 1995.
- [TSA84] R.Y. TSAI und T.S. HUANG: *The perspective view of three points*, IEEE T-PAMI, Bd. 6, S. 13-27, 1984.
- [WAN94] H. WANG und M. BRADY: *A practical solution to corner detection*, Proc. Int. Conf. on Image Processing, Bd. 1, S. 919-923, 1994.
- [WEH98] A. WEHR: *Scannertechniken zur dimensionellen Oberflächenbestimmung*, 44. DVW-Seminar – Hybride Vermessungssysteme, Schriftenreihe des DVW, Bd. 29, S. 125-148, 1998.
- [WIE00] A. WIEDEMANN, M. HEMMLEB und J. ALBERTZ: *Reconstruction of historical buildings based on images from the MEYDENBAUER archives*, IAPRS, Bd. XXXIII, Nr. B5/2, Amsterdam, Niederlande, S. 887-893, 2000.

- [WIL94] R. WILLSON: *Modeling and calibration of automated zoom lenses*, Proc. SPIE 2350, Videometrics III, Boston, USA, S. 170-186, Okt. 1994.
- [ZEL96] C. ZELLER und O.D. FAUGERAS: *Camera self-calibration from video sequences: The KRUPPA equations revisited*, TR 2793, INRIA, Sophia Antipolis, Frankreich, 65 S., Feb. 1996.
- [ZHA98] Z. ZHANG: *Determining the epipolar geometry and its uncertainty – A review*, IJCV, Bd. 27, Nr. 2, S. 161-195, März 1998.

Bei den Literaturreferenzen werden folgende Abkürzungen verwendet:

<b>CAIP</b>	Computer Analysis of Images and Patterns
<b>CAMP</b>	Computer Architectures for Machine Perception
<b>CIPA</b>	Comité International de Photogrammétrie Architecturale
<b>CVIU</b>	Computer Vision and Image Understanding
<b>CVGIP</b>	Computer Vision, Graphics and Image Processing
<b>CVPR</b>	Computer Vision and Pattern Recognition
<b>DAGM</b>	Deutsche Arbeitsgemeinschaft für Mustererkennung
<b>DGPF</b>	Deutsche Gesellschaft für Photogrammetrie und Fernerkundung
<b>DGK</b>	Deutsche Geodätische Kommission
<b>DVW</b>	Deutscher Verein für Vermessungswesen
<b>ECCV</b>	European Conference on Computer Vision
<b>IAPRS</b>	International Archives for Photogrammetry and Remote Sensing
<b>ICCV</b>	International Conference on Computer Vision
<b>ICPR</b>	International Conference on Pattern Recognition
<b>IEEE</b>	Institute of Electrical and Electronic Engineers
<b>IJCV</b>	International Journal of Computer Vision
<b>INRIA</b>	Institut National de Recherche en Informatique et en Automatique
<b>ISPRS</b>	International Society for Photogrammetry and Remote Sensing
<b>LNCS</b>	Lecture Notes in Computer Science
<b>MIT</b>	Massachusetts Institute of Technology
<b>SPIE</b>	International Society for Optical Engineering
<b>T-PAMI</b>	Transactions on Pattern Analysis and Machine Intelligence

**Tab. A.2:** Abkürzungen bei den Literaturangaben

# Formelzeichen

*	Faltungs-Operator
$\ \mathbf{A}\ $	Frobenius-Norm der Matrix $\mathbf{A}$ : $\ \mathbf{A}\  = \sqrt{\sum a_{ij}^2}$
$\mathbf{A}^{-T}$	Transponierte Matrix der Inversen von $\mathbf{A}$ : $(\mathbf{A}^{-1})^T$
$d(\mathbf{x}, \mathbf{y})$	Geometrischer (euklidischer) Abstand der homogenen Koordinaten $\mathbf{x}$ und $\mathbf{y}$
$\det(\mathbf{A})$	Determinante der Matrix $\mathbf{A}$
$\mathbf{D}$	Diagonal-Matrix mit den Elementen $\text{diag}(d_1, \dots, d_n)$
$\exp(x)$	Exponentialfunktion: $e^x$
$\mathbf{e}$	Epipol (3-Vektor)
$\mathbf{E}$	Essentielle Matrix für zwei Bilder ( $3 \times 3$ Rang-2 Matrix)
$f(x, y)$	Bildfunktion
$f_x, f_y$	Erste partielle Ableitung von $f$ in $x$ - bzw. $y$ -Richtung
$f_{xx}, f_{xy}, f_{yy}$	Zweite partielle Ableitungen von $f$
$\mathbf{F}$	Fundamental-Matrix für zwei Bilder ( $3 \times 3$ Rang-2 Matrix)
$\mathcal{F}$	Verschwindungsebene
$G$	GAUSS-Funktion
$\mathbf{H}$	Homographie-Matrix
$\mathbf{I}$	Einheits-Matrix $\text{diag}(1, \dots, 1)$
$\mathbf{K}$	Kalibrierungs-Matrix ( $3 \times 3$ obere Dreiecks-Matrix)
$\mathbf{l}$	Gerade im Bild (3-Vektor)
$\min_x C(x)$	Minimierung der Fehlerfunktion $C$ ( <i>Cost function</i> )
$\nabla$	Gradienten-Operator Nabla
$\boldsymbol{\omega}, \boldsymbol{\omega}^*$	(Duale) Projektion des absoluten Kegelschnitts ( $3 \times 3$ -Matrix)
$\boldsymbol{\Omega}_\infty, \boldsymbol{\Omega}_\infty^*$	(Dualer) absoluter Kegelschnitt ( $3 \times 3$ -Matrix)
$\mathbf{P}$	Projektions-Matrix zur perspektivischen Abbildung ( $3 \times 4$ -Matrix)
$\mathbb{IP}^2, \mathbb{IP}^3$	Projektive Ebene und projektiver Raum
$\Pi_\infty$	Ideale Ebene (4-Vektor)
$\mathbf{Q}_\infty, \mathbf{Q}_\infty^*$	(Duale) absolute Fläche zweiter Ordnung ( $4 \times 4$ Rang-3 Matrix)
$\mathbf{R}$	Rotations-Matrix ( $3 \times 3$ -Orthogonal-Matrix)
$\mathcal{R}$	Projektionsebene ( <i>Retinal plane</i> )
$\sigma, \sigma_i^2, \sigma_{ij}$	Standardabweichung, Varianz und Kovarianz einer GAUSSschen Normalverteilung
$\text{spur}(\mathbf{A})$	Summe der Hauptdiagonalelemente von $\mathbf{A}$ : $\sum a_{ii}$
$\mathbf{T}$	Transformations-Matrix
$\mathbf{x}$	Bildpunkt (3-Vektor)
$\mathbf{X}$	Objektpunkt (4-Vektor)

**Volker Rodehorst**, geboren 1968 in Rustenburg / Südafrika. Nach dem Studium der Informatik wissenschaftlicher Mitarbeiter für Computer Vision sowie Photogrammetrie und Kartographie an der Technischen Universität Berlin.

Forschungsschwerpunkte: Photogrammetrische 3D-Rekonstruktion mittels projektiver Geometrie, wissenschaftliche Bildanalyse in der Architekturphotogrammetrie und Echtzeitalgorithmen für die Stereo-Bildverarbeitung.

**D**ie Photogrammetrie beschäftigt sich mit der Gewinnung geometrischer Objektinformationen durch die Auswertung von Bildern. Dazu müssen die bei der Aufnahme geltenden Abbildungsgesetze rekonstruiert werden. Hier stellt der Nahbereich aufgrund flexibler Bildanordnungen und perspektivischer Verzerrungen eine besondere Herausforderung dar.

In dieser Arbeit wird für die automatische Zuordnung der Bilder eine robuste Merkmalsextraktion und eine verbesserte Korrespondenz-Analyse vorgestellt. Ausgehend von einer algebraischen Darstellung werden schließlich elementare Algorithmen für die projektive Stereo- bzw. Mehrbild-Geometrie behandelt. Unter Berücksichtigung der neuesten Entwicklungen ist auch ein Verfahren zur automatischen Selbst-Kalibrierung integriert.