

## 2 Das Kameramodell

Ein grundlegender Schritt bei der Generierung eines Mosaikbildes besteht in der Registrierung der Einzelbilder einer gegebenen Folge. Dabei werden die zwischen den einzelnen Bildern beobachteten Veränderungen mit Hilfe eines geeigneten Bewegungsmodells  $T_{\vec{p}}$  mathematisch beschrieben. Insbesondere wird in Abhängigkeit von den Daten ein passender Parametersatz  $\vec{p}$  für das Modell geschätzt. Die Bilder können dann durch Anwendung von  $T_{\vec{p}}$  in das gemeinsame Referenzkoordinatensystem transformiert und durch Fusion ihrer Bilddaten zum Mosaikbild zusammengefügt werden. Das Bewegungsmodell selbst leitet sich bei einer bewegten Kamera in der Regel direkt aus der Kamerabewegung ab, wobei seine Komplexität im Wesentlichen mit den Freiheitsgraden der eingesetzten Kamera skaliert. Hieraus resultiert die Notwendigkeit, der Wahl eines geeigneten Bewegungsmodells zunächst eine Modellierung der Kamera voranzustellen.

In diesem Kapitel wird das der Arbeit zu Grunde liegende Modell einer *endlichen, projektiven Kamera* skizziert, aus dem sich direkt das gewählte Bewegungsmodell einer *projektiven Transformation* motiviert. Dabei reichen zur Schätzung von Parametern für dieses Modell grundsätzlich allein die Bilddaten aus, weitere Informationen, d.h. insbesondere Kalibrierungsdaten der Kamera, sind zumeist nicht notwendig. Allerdings erfordert eine Repräsentation der visuellen Daten auf einem polyedrischen Grundkörper, wie sie in der Einleitung als Kernidee der Multi-Mosaikbilder beschrieben wurde, zur korrekten Festlegung der Skalierung des Körpers zumindest die Kenntnis der Bildweite. Daher werden im Unterkapitel 2.4 auch Verfahren zur Online- und Offline-Kalibrierung von endlichen, projektiven Kameras diskutiert. Die Darstellung in diesem Kapitel orientiert sich im Wesentlichen an [Har00], wobei zur Formulierung mathematischer Sachverhalte unter anderem homogene Koordinaten aus der projektiven Geometrie verwendet werden. Aus diesem Grund findet sich im nachfolgenden Abschnitt zunächst eine kurze Einführung in die projektive Geometrie, einschließlich der verwendeten Notation. Weiterführende Details sind in [Moh96] oder [Har00] zu finden.

### 2.1 Grundzüge der projektiven Geometrie

Die projektive Geometrie hat nach [Edg02] ihren Ursprung in der Malerei der italienischen Renaissance. Dort erwuchs das Bestreben, Eindrücke aus der 3D-Welt, insbesondere bezogen auf die Architektur, korrekt in Gemälden widerzuspiegeln und in diesem Zusammenhang mathematisch zu formalisieren. Die gängige Geometrie der euklidischen Räume  $\mathbb{R}^n$  stieß dabei an ihre Grenzen. Insbesondere projektive Effekte im Unendlichen, wie beispielsweise der Schnitt zweier parallel verlaufender Geraden im imaginären

Fluchtpunkt unter projektiver Abbildung auf eine Ebene, erforderten neue Formalismen, die in der Einführung projektiver Räume  $\mathbb{P}^n$  gipfelten. Diese stellen im Wesentlichen eine Erweiterung der euklidischen Räume  $\mathbb{R}^n$  um *ideale* Punkte, Geraden und Ebenen dar, die im Unendlichen liegen und damit den Ausgangspunkt zur mathematischen Beschreibung entsprechender Effekte bilden.

Die Grundlage projektiver Räume ist durch *homogene Koordinaten* gegeben, deren Haupteigenschaft die Invarianz gegenüber Skalierungen ist. Punkte im  $\mathbb{P}^n$  werden durch  $(n+1)$ -dimensionale Vektoren beschrieben, wobei Nullvektoren  $\vec{0}$  ausgeschlossen sind. Die Skalierungsinvarianz lässt sich durch eine Äquivalenzrelation ausdrücken, die nachfolgend exemplarisch für Punkte  $u$  des  $\mathbb{P}^2$  angegeben ist:

$$u = (x, y, z)^T \sim \lambda (x, y, z)^T, \quad \lambda \neq 0, \lambda \in \mathbb{R}.$$

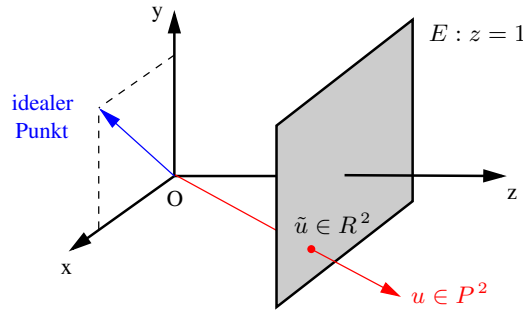
Ein aus dieser Eigenschaft leicht abzuleitendes Modell für den  $\mathbb{P}^2$ , das zum besseren Verständnis projektiver Räume hier vorgestellt werden soll, beschreibt die projektiven Punkte des Raumes als Strahlen, die durch den Ursprung eines 3D-Koordinatensystems verlaufen (Abb. 2.1). Alle auf einem Strahl liegenden 3D-Punkte sind durch Multiplikation mit einem geeigneten Skalierungsfaktor ineinander überführbar und somit in diesem Modell äquivalent. Werden die Strahlen des  $\mathbb{P}^2$  mit einer beliebigen Ebene des Raumes geschnitten, so lassen sich die resultierenden Schnittpunkte als eindeutige Repräsentanten der Strahlen auffassen. Dabei existieren für parallel zur jeweils gewählten Ebene verlaufende Strahlen keine Repräsentanten. Im Sonderfall der parallel zur xy-Ebene gelegenen Ebene  $E : z = 1$  ergeben sich alle Repräsentanten aus den geschnittenen Strahlen durch Skalierung der dritten Vektorkomponenten auf 1. Vernachlässigt man darüber hinaus diese dritten Komponenten der Repräsentanten vollständig, resultiert gerade die Menge der zweidimensionalen Vektoren des  $\mathbb{R}^2$ . Die Ebene  $E : z = 1$  lässt sich somit als Einbettung des  $\mathbb{R}^2$  in den  $\mathbb{P}^2$  auffassen, wobei der  $\mathbb{P}^2$  gegenüber dem  $\mathbb{R}^2$  noch um die Punkte ergänzt wird, deren Strahlen die Ebene nicht schneiden. Letztere verlaufen parallel zu  $E$  in der xy-Ebene des 3D-Koordinatensystems und weisen in ihrer dritten Komponente eine Null auf. Sie liegen im Prinzip im Unendlichen und werden als *ideale* Punkte bezeichnet. Formal lässt sich zu jedem nicht-idealen Punkt  $u \in \mathbb{P}^2$  ein eindeutiger 2D-Repräsentant  $\tilde{u} = (x, y)^T \in \mathbb{R}^2$  durch folgende Funktion bestimmen:

$$f_{PR} : u = (x, y, z)^T \rightarrow \tilde{u} = (x/z, y/z)^T \quad \text{mit} \quad (x, y, z)^T \sim (x/z, y/z, 1)^T. \quad (2.1)$$

Umgekehrt kann für jeden Punkt des  $\mathbb{R}^2$  ein Vertreter der repräsentierten Äquivalenzklasse im  $\mathbb{P}^2$  durch Ergänzung einer 1 in der dritten Vektorkomponente ermittelt werden:

$$f_{RP} : \tilde{u} = (x, y)^T \rightarrow u = (x, y, 1)^T. \quad (2.2)$$

Im Rahmen dieser Arbeit werden zur Bezeichnung von Punkten des  $\mathbb{R}^2$  bzw.  $\mathbb{P}^2$  Kleinbuchstaben verwendet, und für Punkte aus  $\mathbb{R}^3$  und  $\mathbb{P}^3$  Großbuchstaben. Punkte aus den euklidischen Räumen werden zur Unterscheidung zusätzlich mit einer Tilde versehen.



**Abbildung 2.1:** Visualisierung des Strahlenmodells des  $\mathbb{P}^2$ , wobei die eingezeichnete Ebene  $E: z = 1$  die Einbettung des  $\mathbb{R}^2$  in den  $\mathbb{P}^2$  veranschaulicht.

Neben dem mathematischen Formalismus zur Beschreibung projektiver Effekte bringen homogene Koordinaten einen weiteren Vorteil auch im Hinblick auf andere Anwendungsfelder mit sich. Viele mathematische Konzepte, insbesondere Abbildungen, lassen sich unter Verwendung von homogenen Koordinaten durch leicht zu handhabende Matrix-Vektor-Operationen realisieren. Das nachfolgende Beispiel veranschaulicht dies. Eine affine Transformation  $A$  im  $\mathbb{R}^2$  setzt sich aus einer Rotationsmatrix  $R$  und einem Translationsvektor  $\vec{t}$  zusammen:

$$A(\tilde{u}) = R \cdot \tilde{u} + \vec{t} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \cdot \tilde{u} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}.$$

In homogenen Koordinaten lässt sich die obige Gleichung durch eine einfache Matrixmultiplikation ausdrücken:

$$A(u) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & t_x \\ \sin(\alpha) & \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot u = \begin{bmatrix} R & \vec{t} \\ \vec{0}^T & 1 \end{bmatrix} \cdot u.$$

Die Darstellung der Matrix im hinteren Teil der Gleichung, die Teilmengen der Matrixkomponenten in *Blöcken* zusammenfasst, wird als *Block-Schreibweise* bezeichnet und bietet sich an, wenn die einzelnen Komponenten der Matrix nicht direkt von Interesse sind. Dabei können die einzelnen Blöcke zur besseren Lesbarkeit durch senkrechte Striche voneinander abgegrenzt werden (vgl. etwa Gl. 2.3).

## 2.2 Endliche, projektive Kameras

Der im Rahmen dieser Arbeit entwickelte, visuelle Speicher dient der Repräsentation von Bilddaten, die mit stationären, jedoch um ihr optisches Zentrum rotierenden und zoomenden Kameras aufgenommen werden. Für den Aufbau der Kameras wird dabei das *Lochkameramodell* zu Grunde gelegt (Abb. 2.2). Das Projektionszentrum der Kamera  $O_{cam}$  (*optisches Zentrum*) liegt im Ursprung eines 3D-Kamerakoordinatensystems, die optische Achse fällt mit der z-Achse des Koordinatensystems zusammen. Im Abstand der

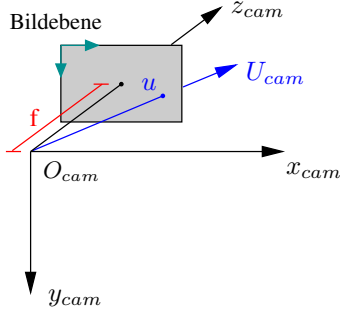


Abbildung 2.2: Das Lochkameramodell.

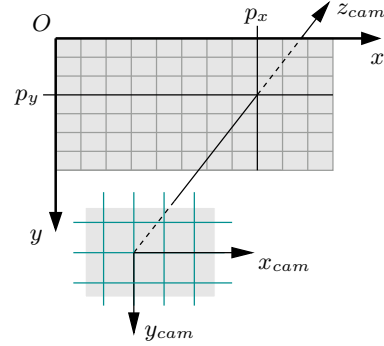


Abbildung 2.3: Bildkoordinatensystem.

*Bildweite (focal length)*  $f > 0$  vom optischen Zentrum befindet sich die Bildebene, auf die die betrachtete 3D-Welt zweidimensional abgebildet wird. Die Bildebene ist dabei stets parallel zur  $xy$ -Ebene des 3D-Koordinatensystems ausgerichtet. Sie besitzt ein eigenes lokales 2D-Koordinatensystem (Abb. 2.3), dessen Ursprung in der Regel in die linke obere Ecke der Ebene gelegt wird (*upper left*).

Die Position eines Punktes auf dieser Ebene wird in homogenen Koordinaten angegeben, die sich aus der Anwendung der Funktion  $f_{RP}$  aus Gleichung 2.2 auf die lokalen Bildkoordinaten ergeben. Der *Hauptpunkt* der Kamera, an dem die optische Achse die Bildebene durchstößt, liegt entsprechend an Position  $(p_x, p_y, 1)^T$  auf der Ebene.

Die Projektion  $u$  eines beliebigen 3D-Punktes  $U_{cam}$  des Kamerakoordinatensystems auf die Bildebene ergibt sich in diesem Modell als Schnittpunkt des Richtungsstrahls vom optischen Kamerazentrum zum Punkt  $U_{cam}$  mit der Bildebene. Er kann unter Anwendung des Strahlensatzes in homogenen Koordinaten wie folgt berechnet werden:

$$u = K [I | \vec{0}] \cdot U_{cam} \quad \text{mit} \quad K = \begin{bmatrix} f & p_x \\ f & p_y \\ 1 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.3)$$

Dabei ist noch unberücksichtigt geblieben, dass die beiden Koordinatenachsen der Bildebene in der Regel jeweils unterschiedlich und darüber hinaus unabhängig vom 3D-Koordinatensystem der Kamera skaliert sind. Auf diese Weise wird der Tatsache Rechnung getragen, dass die einzelnen Zellen der in den Kameras üblicherweise zur Bildaufnahme eingesetzten CCD-Arrays<sup>1</sup> nicht immer quadratisch sind. Um die Bildkoordinaten in Pixeleinheiten korrekt errechnen zu können, müssen somit bei der Abbildung in  $x$ - und  $y$ -Richtung jeweils spezifische Skalierungsfaktoren berücksichtigt werden. Wenn  $m_x$  und  $m_y$  die Anzahlen an CCD-Elementen pro Einheitslänge im Kamerakoordinatensystem in beiden Achsenrichtungen angeben, so sind die 3D-Koordinaten bei der Abbildung mit den Faktoren  $\alpha_x = f \cdot m_x$  und  $\alpha_y = f \cdot m_y$  zu multiplizieren, gegeben die Bildweite  $f$  in Pixeln. Ergänzend ist auch der Hauptpunkt umzurechnen.

<sup>1</sup>CCD-Arrays sind hier und im weiteren Verlauf der Arbeit als Synonym für pixelbasierte Aufnahmetechniken zu verstehen, zu denen beispielsweise auch CMOS-Verfahren gehören.

Zusätzlich wird zur Quantifizierung von Abweichungen der Achsen des Bildkoordinatensystems von der üblicherweise angenommenen Orthogonalität noch ein weiterer Parameter  $s$  (*Skew*) eingeführt. Für normale Kameras und Abbildungen wird er aber im Allgemeinen zu Null angenommen. Die vollständige *Kalibrierungsmatrix*  $K$  aus Gleichung 2.3 ergibt sich folglich unter Berücksichtigung dieser Aspekte zu

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \quad \text{mit} \quad x_0 = p_x \cdot m_x, \quad y_0 = p_y \cdot m_y. \quad (2.4)$$

Insgesamt werden somit bislang zur Beschreibung des Kameramodells und damit auch der projektiven Abbildung die fünf *intrinsischen* Parameter  $x_0, y_0, \alpha_x, \alpha_y$  und  $s$  benötigt. Die Kamera ist jedoch in ein 3D-Weltkoordinatensystem eingebettet, das ihre globale Position in der Szene festlegt. In diesem Fall lassen sich 3D-Punkte in der Welt nur dann direkt unter Anwendung von Gleichung 2.3 auf die Bildebene projizieren, wenn Kamera- und Weltkoordinatensystem übereinstimmen. Andernfalls muss die Transformation zwischen den beiden Systemen explizit in die Abbildung einbezogen werden.

Zwei gleichartig orientierte<sup>2</sup> 3D-Koordinatensysteme lassen sich durch Rotation und Translation ineinander überführen. Für einen Punkt  $U$  im 3D-Weltkoordinatensystem bedeutet dies, dass er vor der eigentlichen Abbildung in 3D-Kamerakordinaten  $U_{cam}$  umgerechnet werden muss:

$$U_{cam} = \begin{bmatrix} R & -R\tilde{C} \\ \vec{0}^T & 1 \end{bmatrix} \cdot U.$$

In der vorstehenden Formel bezeichnet  $R$  die  $3 \times 3$ -dimensionale Rotationsmatrix zwischen den beiden Koordinatensystemen, die durch drei Drehwinkel um die einzelnen Koordinatenachsen parametrisiert ist. Der 3D-Vektor  $\tilde{C}$  gibt das Zentrum der Kamera in euklidischen Weltkoordinaten an und kodiert damit die durchzuführende Translation.  $R$  und  $\tilde{C}$  werden unter dem Oberbegriff der *extrinsischen* Kameraparameter zusammengefasst. Für das Gesamtmodell der Kamera ergeben sich unter Berücksichtigung der intrinsischen und extrinsischen Parameter insgesamt elf Freiheitsgrade, die in der *Kamera- oder Abbildungsmatrix*  $P$  zusammengefasst werden:

$$P = K [I | \vec{0}] \cdot \begin{bmatrix} R & -R\tilde{C} \\ \vec{0}^T & 1 \end{bmatrix} = K R [I | -\tilde{C}] = [M | -M\tilde{C}]. \quad (2.5)$$

Ein solches Kameramodell wird als *endliche, projektive Kamera* bezeichnet, wobei die Nicht-Singularität der Matrix  $M = KR$  entscheidend ist. Die Abbildung eines Punktes  $U$  auf die Bildebene ist abschließend durch folgende Formel gegeben:

$$u = K [I | \vec{0}] \cdot U_{cam} = P \cdot U = K R [I | -\tilde{C}] \cdot U. \quad (2.6)$$

<sup>2</sup>3D-Koordinatensysteme können entweder rechts- oder linkshändig orientiert sein.

## 2.3 Projektive Transformationen

Bei der Generierung eines Mosaikbildes müssen einzelne Bilder einer Folge zueinander registriert werden. Dies bedeutet insbesondere, dass Zusammenhänge zwischen den projektiven Abbildungen einer Szene auf verschiedene Bildebenen herzustellen sind. Während dies im allgemeinen Fall beliebiger Kameras und Konstellationen sehr komplex sein kann, vereinfacht sich die Situation, wenn endliche, projektive Kameras zu Grunde gelegt werden, die ortsgebunden sind und lediglich um ihre optischen Zentren rotieren (Abb. 2.4). In diesem Fall kann über die Gleichungen 2.5 und 2.6 direkt ein Zusammenhang zwischen den Abbildungsmatrizen hergeleitet werden. Sei  $P = KR[I | -\tilde{C}]$  die Kameramatrix der ersten Kamera, und  $P' = K'R'[I | -\tilde{C}]$  die der zweiten. Dann resultiert aus einer Gleichsetzung der identischen Teilmatrizen  $[I | -\tilde{C}]$  direkt der folgende Zusammenhang:

$$P = (KR) \cdot (K'R')^{-1} \cdot P'.$$

Für die Abbildungen  $u$  und  $u'$  eines Punktes  $U$  aus der Szene auf die zwei verschiedenen Bildebenen gilt damit:

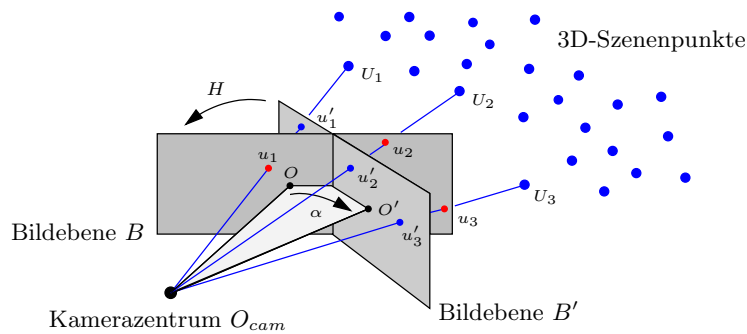
$$u = P \cdot U = (KR) \cdot (K'R')^{-1} \cdot P' \cdot U = (KR) \cdot (K'R')^{-1} \cdot u'. \quad (2.7)$$

Basierend auf der bei endlichen, projektiven Kameras gegebenen Nicht-Singularität der Matrizen  $KR$  und  $K'R'$  lässt sich das Matrixprodukt aus der vorstehenden Gleichung 2.7 schließlich durch eine einzelne, nicht-singuläre  $3 \times 3$ -Matrix  $H$  ersetzen, die eine Abbildung vom  $\mathbb{P}^2$  in den  $\mathbb{P}^2$  realisiert:

$$u = (KR) \cdot (K'R')^{-1} \cdot u' = H \cdot u' \quad , \quad H : \mathbb{P}^2 \rightarrow \mathbb{P}^2. \quad (2.8)$$

Nicht-singuläre  $3 \times 3$ -Matrizen mit dieser Eigenschaft werden allgemein als *projektive Transformationen* oder *Homographien* bezeichnet. Sie sind aufgrund der Skalierungsinvarianz homogener Koordinaten ebenfalls skalierungsinvariant, d.h. sie lassen sich durch acht Parameter exakt spezifizieren. In der praktischen Anwendung von Homographien werden die neun Einträge der Homographiematrix daher in der Regel konsistent skaliert, wobei beispielsweise der Eintrag rechts unten in der Matrix auf 1 normiert werden kann. Bei der Abbildung von Punktemengen aus dem  $\mathbb{P}^2$  mit Homographien bleiben gerade Linien grundsätzlich als solche erhalten, Längenverhältnisse und Winkel können sich jedoch verändern. Die Menge der Homographien, die der Menge aller reellen, regulären Matrizen entspricht, bildet bezüglich der Matrixmultiplikation eine Gruppe. Daraus folgt, dass sich jede projektive Transformation invertieren lässt. Außerdem wird dadurch auch die Konkatenation 'o' von Homographien eindeutig definiert, wobei sie auf die Multiplikation der Homographiematrizen zurückgeführt wird.

Aus den vorstehenden Ausführungen folgt, dass sich Bilder, die mit einer ortsfesten, rotierenden Kamera bei variierenden intrinsischen Parametern aufgenommen wurden, durch die Berechnung einer Homographie zueinander registrieren lassen. Dieses Modell bildet damit die Grundlage zur Bildregistrierung in der vorliegenden Arbeit. Dabei sei



**Abbildung 2.4:** Modell einer stationären, rotierenden Kamera: Zwischen den Abbildungen von 3D-Punkten auf die beiden Bildebenen  $B$  und  $B'$  besteht ein projektiver Zusammenhang  $H$ .

insbesondere darauf hingewiesen, dass zur reinen Verknüpfung der Bilder weder intrinsische noch extrinsische Kameraparameter bekannt sein müssen. Diese werden erst für eine metrische Rekonstruktion der betrachteten Szene und im Kontext dieser Arbeit vorrangig zur Auswahl geeigneter Projektionskörper für die Multi-Mosaikbilder benötigt. Aus diesem Grund werden im nachfolgenden Abschnitt Methoden zur Bestimmung der Kameraparameter (*Kalibrierung*) vorgestellt und diskutiert.

## 2.4 Kamerakalibrierung

Das Ziel einer *Kalibrierung* besteht in der Rekonstruktion der extrinsischen und intrinsischen Kameraparameter, mit denen vorliegende Bilder aufgenommen wurden. Projektive Zusammenhänge zwischen Bildern im Rahmen einer Berechnung von Mosaikbildern lassen sich auch ohne Kenntnis der konkreten Parameterwerte schätzen. Für eine Extraktion von Tiefen- oder 3D-Informationen einer betrachteten Szene sind diese Daten jedoch oftmals unerlässlich. Obgleich die vorliegende Arbeit nicht auf eine solche 3D-Szenenrekonstruktion zielt, kann aber auch hier nicht auf eine Kalibrierung der Kameras verzichtet werden. Dies resultiert aus der Tatsache, dass eine möglichst verzerrungsfreie Projektion von Bilddaten auf Polyeder, wie sie das in der Einleitung skizzierte Konzept des visuellen Speichers vorsieht, eine korrekte Skalierung der polyedrischen Grundkörper voraussetzt, und diese direkt von der Bildweite  $f$  der Kameras abhängt.

In der Literatur existieren verschiedene Ansätze zur Bestimmung der intrinsischen und extrinsischen Kameraparameter  $K$ ,  $R$  und  $\tilde{C}$ . Im Allgemeinen erfolgt dabei keine explizite Rekonstruktion der Bildweite  $f$ , sondern die Skalierungsfaktoren  $\alpha_x$  und  $\alpha_y$  aus Gleichung 2.4 werden direkt als achsenspezifische Bildweiten interpretiert.  $f$  lässt sich zwar mit Hilfe der konkreten Werte für  $m_x$  bzw.  $m_y$ , die der technischen Spezifikation einer Kamera entnommen werden können, auch exakt berechnen,  $\alpha_x$  und  $\alpha_y$  spezifizieren die gesuchte Abbildung jedoch zumeist bereits hinreichend eindeutig, so dass üblicherweise auf eine explizite Bestimmung von  $f$  verzichtet wird.

Die Grundidee vieler Verfahren besteht darin, die Parameter aus Paaren von 3D-Szenenpunkten mit bekannten Weltkoordinaten und ihren 2D-Abbildungen in aufgenomme-



**Abbildung 2.5:** Eingesetzte Kameratypen der Firma Sony<sup>TM</sup>: links die zwei Kameras des Typs „DFW-VL500“, montiert auf dem Stereokamerakopf des Typs „PowerCube“ der Firma Amtec<sup>TM</sup>, rechts die Kamera des Typs „EVI-D31“. Die Pfeile kennzeichnen jeweils verfügbare Rotationsachsen.

nen Bildern zu rekonstruieren. Die Festlegung des 3D-Weltkoordinatensystems erfolgt dabei zumeist anhand eines geeigneten Kalibrieramusters (z.B. Abb. 2.6(a)), so dass diese Art der Kalibrierung *offline* durchgeführt werden muss. Als Alternative, die auch *online* angewandt werden kann, haben sich in der Vergangenheit *Autokalibrierverfahren* etabliert. Insbesondere im Rahmen von Homographieschätzungen zwischen Bildern wurden verschiedene Algorithmen publiziert, die eine Rekonstruktion der intrinsischen Parameter  $K$  erlauben. In den beiden folgenden Abschnitten werden ausgewählte Ansätze zur Offline- und Online-Kalibrierung vorgestellt, während Abschnitt 2.4.1 Ergebnisse einer Kalibrierung der in dieser Arbeit eingesetzten Kameras (Abb. 2.5) enthält.

### Offline-Kalibrierung

Im Allgemeinen stellt eine Menge von Paaren dreidimensionaler Szenenpunkte und ihrer korrespondierenden zweidimensionalen Abbildungen in aufgenommenen Bildern den Ausgangspunkt einer Offline-Kalibrierung zur Schätzung der intrinsischen und extrinsischen Parameter einer Kamera dar. Das mithin gängigste Verfahren auf dieser Basis zur direkten Ermittlung der intrinsischen und extrinsischen Parameter, sowie zusätzlich von Korrekturfaktoren für eine eventuell vorliegende Linsenverzerrung, wurde bereits 1986 von Tsai veröffentlicht [Tsa86]. Der Algorithmus arbeitet in zwei Schritten, wobei zunächst für möglichst viele Parameter initiale Schätzungen über einen effizienten, linearen Optimierungsansatz ermittelt werden. Anschließend erfolgt im Rahmen einer iterativen, nichtlinearen Optimierung eine Verfeinerung dieser initialen Werte, wobei auch Nebenbedingungen für die einzelnen Parameter in die Berechnungen einfließen.

Eine alternative Methode, die von Hartley [Har00] publiziert wurde, basiert im Gegensatz dazu auf einer expliziten Berechnung der Kameramatrix  $P$  (Gl. 2.5). Sie lässt sich bestimmen, indem die gegebenen 2D/3D-Punktkorrespondenzen in die Abbildungsgleichung projektiver Kameras (Gl. 2.6) eingesetzt werden. Die resultierenden Bedingungsgleichungen können zu einem (in der Regel überbestimmten) linearen Gleichungssystem zusammengefasst werden, das beispielsweise über eine Singulärwertzerlegung lösbar ist. Dabei sind für eine vollständige Schätzung aller elf Parameter mindestens sechs Punktkorrespondenzen bei zwei Bedingungsgleichungen pro Korrespondenz erforderlich. Aus der auf diese Weise ermittelten Matrix  $P$  können abschließend die gesuchten Vektoren



und Matrizen  $\tilde{C}$ ,  $R$  und  $K$  hergeleitet werden ([Har00], S. 150ff.).  $\tilde{C}$  resultiert aus einer Singulärwertzerlegung von  $P$ , während  $R$  und  $K$  mit Hilfe einer  $RQ$ -Zerlegung aus  $M$  abgeleitet werden können (s. Gl. 2.5). Die  $RQ$ -Zerlegung von  $M$  errechnet sich dabei, indem  $M^T$  zunächst mit einer Hilfsmatrix  $T$  multipliziert wird. Das Ergebnis dieser Operation kann dann mit einer herkömmlichen  $QR$ -Zerlegung ([Pre92], S. 98ff.) in zwei Matrizen  $\bar{Q}$  und  $\bar{R}$  aufgespalten werden, aus denen sich schließlich die gesuchten Matrizen  $R$  und  $Q$  der  $RQ$ -Zerlegung von  $M$  bestimmen lassen:

$$M^T \cdot T = M^T \cdot \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \bar{Q}\bar{R} \quad , \quad \text{so dass} \quad R = T\bar{R}^T T \quad \text{und} \quad Q = T\bar{Q}^T.$$

Die Qualität der aus einer Offline-Kalibrierung resultierenden Kameraparameter hängt, unabhängig von dem eingesetzten Kalibrierverfahren, direkt mit der Genauigkeit der 3D-Welt- und der 2D-Bildkoordinaten korrespondierender Punkte zusammen. Folglich sollte ein möglichst exaktes Muster zur Kalibrierung verwendet werden, das eine einfache und vor allem auch exakte Lokalisation ausgewählter Punkte in den Bildern bis auf Subpixelgenauigkeit unterstützt. Da die Qualität der Kalibrierung weiterhin mit der Anzahl zu Grunde gelegter Punktpaare zunimmt, ist eine automatische Punktdetektion wünschenswert, um eine mit einem sehr hohen Aufwand verbundene und zudem relativ ungenaue manuelle Auswahl von Punkten vermeiden zu können.

## Online-Autokalibrierung

Eine Kamerakalibrierung mit Hilfe von Kalibriermustern wird im Allgemeinen vor Beginn des eigentlich implizierten Einsatzzieles der Kameras durchgeführt. Dabei sind oftmals für verschiedene Einstellungen einer Kamera voneinander unabhängige Kalibrierungen erforderlich. Eine solche Vorgehensweise ist somit aufwändig, beschränkt die Flexibilität beim Einsatz von Kameras in Anwendungsfeldern, in denen eine Kalibrierung unumgänglich ist, und erzwingt die Bereitstellung eines möglichst exakten Kalibriermusters. *Autokalibrierverfahren* umgehen diese Beschränkungen und erlauben eine automatische, schritthaltende Kamerakalibrierung, die beispielsweise parallel zu einer Online-Generierung von Mosaikbildern erfolgen kann. Die Grundidee der Verfahren beruht auf einer Rekonstruktion der Parameter aus Korrespondenzen zwischen *verschiedenen* Bildern, wobei *keine* bekannten 3D-Weltkoordinaten mehr vorausgesetzt werden. In Abhängigkeit von den bei der Bildaufnahme ausgeführten Kamerabewegungen lassen sich dabei verschiedene Teilmengen der Parameter ermitteln [Fau92, Har99, Pol99]. Im Kontext der vorliegenden Arbeit sind vorrangig die intrinsischen Parameter in  $K$  von Interesse, die unmittelbar auf Basis geschätzter Homographien bestimmt werden können.

Shum und Szeliski propagieren dazu in [Shu00] eine direkte Berechnung der Bildweite  $f$  aus den Einträgen einer Homographiematrix (vgl. S. 31):

$$f^2 = \frac{h_{23}^2 - h_{13}^2}{h_{11}^2 + h_{12}^2 - h_{21}^2 - h_{22}^2} \quad \text{oder alternativ} \quad f^2 = -\frac{h_{13} \cdot h_{23}}{h_{11} \cdot h_{21} + h_{12} \cdot h_{22}}.$$

Dabei werden quadratische CCD-Elemente ( $\alpha_x = \alpha_y$ ) und eine Positionierung des Hauptpunktes im Bildzentrum (d.h.  $x_0 = y_0 = 0$ , für ein Bildkoordinatensystem, dessen Ursprung im Hauptpunkt liegt) vorausgesetzt.

Ein zweites Verfahren, das eine Kombination verschiedener Bedingungen für die zu schätzenden Kameraparameter erlaubt, stammt von de Agapito und Kollegen [dA99]. Der lineare Ansatz beruht auf den Eigenschaften des *absolute Conic*  $\Omega_\infty$  des  $\mathbb{P}^3$ , dessen Punktmenge durch die folgende Bedingungsgleichung definiert wird ([Har00], S. 63ff.):

$$\Omega_\infty : x_1^2 + x_2^2 + x_3^2 = 0 \wedge x_4 = 0.$$

Es ist direkt ersichtlich, dass der Conic ausschließlich aus idealen Punkten besteht. Seine Abbildung  ${}^i\omega$  durch eine endliche, projektive Kamera  $P_i$  hängt allein von deren intrinsischen Parametern  $K_i$  ab (unter der Bedingung  $s = 0$ ):

$${}^i\omega = K_i^{-T} \cdot K_i^{-1} = \begin{bmatrix} 1/\alpha_x^2 & 0 & -x_0/\alpha_x^2 \\ 0 & 1/\alpha_y^2 & -y_0/\alpha_y^2 \\ -x_0/\alpha_x^2 & -y_0/\alpha_y^2 & 1 + x_0^2/\alpha_x^2 + y_0^2/\alpha_y^2 \end{bmatrix} = \begin{bmatrix} {}^i\omega_{11} & {}^i\omega_{12} & {}^i\omega_{13} \\ {}^i\omega_{12} & {}^i\omega_{22} & {}^i\omega_{23} \\ {}^i\omega_{13} & {}^i\omega_{23} & {}^i\omega_{33} \end{bmatrix}.$$

Mit Hilfe dieser Definition und Gleichung 2.7 lässt sich ein Zusammenhang zwischen den Abbildungen  ${}^i\omega$  und  ${}^j\omega$  des Conic auf zwei verschiedene Bildebenen  $B_i$  und  $B_j$  und der zwischen den Ebenen gegebenen, projektiven Abbildung  $H_{ij}$  herstellen:

$$H_{ij} = K_j R_j R_i^{-1} K_i^{-1} \iff H_{ij} = K_j R_{ij} K_i^{-1} \iff R_{ij} = K_j^{-1} H_{ij} K_i.$$

Da  $R_{ij}$  eine Rotationsmatrix ist, gilt  $R_{ij} R_{ij}^T = I$ . Setzt man die Definitionen aus den beiden vorstehenden Gleichungen in diese Bedingung ein, folgt daraus durch Umformung:

$$H_{ij}^{-T} K_i^{-T} K_i^{-1} H_{ij}^{-1} = K_j^{-T} K_j^{-1} \iff H_{ij}^{-T} {}^i\omega H_{ij}^{-1} = {}^j\omega.$$

Für die einzelnen Komponenten der Abbildung  ${}^j\omega$  von  $\Omega_\infty$  auf die Bildebene  $B_j$  können nun Bedingungen formuliert werden, die sich aus den Annahmen über die Kameraparameter ableiten. Aus der Voraussetzung eines verschwindenden Skews folgt direkt  ${}^j\omega_{12} = 0$ . Liegen darüber hinaus quadratische CCD-Elemente vor, gilt  $\alpha_x = \alpha_y$  und damit  ${}^j\omega_{11} = {}^j\omega_{22}$ . Schlussendlich führt ein im Mittelpunkt der Bilder gelegener Hauptpunkt mit  $x_0 = y_0 = 0$  auf die beiden weiteren Bedingungen  ${}^j\omega_{13} = 0$  und  ${}^j\omega_{23} = 0$ . Alle Bedingungsgleichungen können zu einem linearen Gleichungssystem in den unbekannten Einträgen aus  ${}^i\omega$  kombiniert werden, dessen Lösung schließlich auch die Bestimmung der intrinsischen Parameter  $K_i$  erlaubt. In Abhängigkeit von den zu Grunde gelegten Bedingungen sind dabei zwei, drei oder fünf Homographien erforderlich, um die notwendige Mindestanzahl von fünf Bedingungsgleichungen zur Berechnung der fünf gesuchten Parameter aus  $K_i$  formulieren zu können.

Grundsätzlich ist bei diesem Verfahren zu berücksichtigen, dass nicht alle Kamerabewegungen einer rotierenden und zoomenden Kamera eine stabile Kalibrierung erlauben und die Qualität der Ergebnisse darüber hinaus eng mit den verwendeten Annahmen verzahnt ist (siehe z.B. [Stu97, dA00, Wan04] bzw. den nachfolgenden Abschnitt).

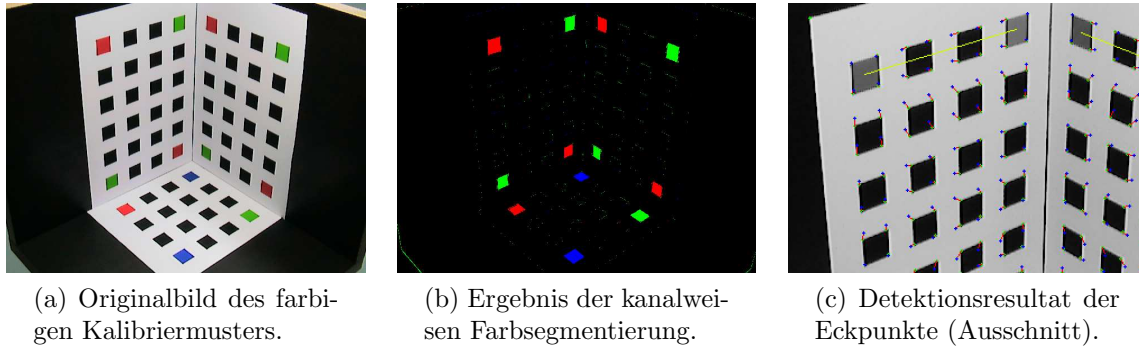
### 2.4.1 Ergebnisse & Auswertung

Im Rahmen dieser Arbeit kommen zwei Kamertypen der Firma Sony<sup>TM</sup> zum Einsatz (Abb. 2.5). Während die Kamera des Typs „EVI-D31“ direkt angesteuert und in horizontaler ( $\approx \pm 100^\circ$ ) und vertikaler ( $\approx \pm 25^\circ$ ) Richtung um das optische Zentrum rotiert und gezoomt werden kann, erlauben die beiden Kameras des Typs „DFW-VL500“ lediglich eine aktive Änderung des Zooms und keine direkte Rotation. Sie lassen sich jedoch auf einem ansteuerbaren Stereo-Kamerakopf der Firma Amtec<sup>TM</sup> montieren, womit eine Bewegung der Kameras ermöglicht wird. Der Kopf weist dabei insgesamt vier Dreh- bzw. Kippgelenke auf, von denen allerdings nur die oberen beiden eine Drehung der Kameras um die optischen Zentren gewährleisten (s. Abb. 2.5). Somit sind lediglich horizontale Rotationen durchführbar, die jedoch einen Drehwinkel von  $360^\circ$  überschreiten können. Die Kamera „EVI-D31“ verfügt zudem über einen Autofokusmechanismus, während sich die beiden anderen Kameras ausschließlich manuell fokussieren lassen.

Beide Kamertypen wurden zunächst unter Anwendung des von Hartley publizierten Verfahrens offline kalibriert, um einen funktionalen Zusammenhang zwischen den Hardwareeinstellungen der Kameras und den jeweils korrespondierenden, intrinsischen Parametern herzustellen. Mit Hilfe dieser Funktionen lassen sich anschließend im Verlauf der Mosaikbildberechnung die jeweils aktuellen Hardware-Zoomeinstellungen direkt auf die korrespondierenden Bildweiten abbilden. Dem System stehen somit bei dieser Vorgehensweise insbesondere auch über längere Zeiträume hinweg konsistente, in ihrer Genauigkeit nicht variierende Werte für die Bildweite zur Verfügung.

Der Zoomparameter der Kamera des Typs „EVI-D31“ weist einen Einstellungsbereich von 0 bis 1023 auf, während die Parameter der beiden anderen Kameras Einstellungen zwischen 40 und 1400 zulassen. Im Rahmen der Kalibrierung wurden für beide Kamertypen jeweils Bildfolgen des fokussierten Kalibriermusters aufgenommen, wobei eine schrittweise Erhöhung der Zoomparameter erfolgte. Dabei ist mit jeder der kalibrierten Zoomeinstellungen implizit ein spezifischer Fokusparameter verknüpft, so dass mögliche Auswirkungen einer Fokusänderung auf die Bildweite (bei konstantem Zoomparameter) zunächst unberücksichtigt bleiben.

Die Detektion der Punktreferenzen für die Kalibrierung erfolgt automatisch. Das verwendete Kalibriermuster (Abb. 2.6(a)) zeigt 52 schwarze und 12 farbkodierte Quadrate mit einer Kantenlänge und einem paarweisen Abstand von jeweils  $2\text{ cm}$ , wobei die geschätzte Genauigkeit in den 3D-Positionen bei etwa  $\pm 0,5\text{ mm}$  liegt. Die insgesamt 256 Eckpunkte lassen sich mit einem Harris-Corner-Detektor [Har88] subpixelgenau detektieren. Ihre anschließende, eindeutige Zuordnung zu den durch das Kalibriermuster vorgegebenen 3D-Punktkoordinaten erfolgt anhand von äquidistant eingeteilten Rastern, die mit Hilfe der farbigen Quadrate automatisch generiert werden und die erwarteten 2D-Positionen der Eckpunkte aller Quadrate auf den einzelnen Ebenen vorgeben. Dazu findet zunächst eine Lokalisation der Farbquadrate in den Bildern durch eine bildspezifische Schwellwertbildung in den jeweiligen Farbkanälen statt (Abb. 2.6(b)), gefolgt von einer Regionensegmentierung. Die Anordnung der Quadrate auf dem Kalibriermuster ermöglicht dabei eine eindeutige Zuordnung der detektierten Regionen zu den einzelnen

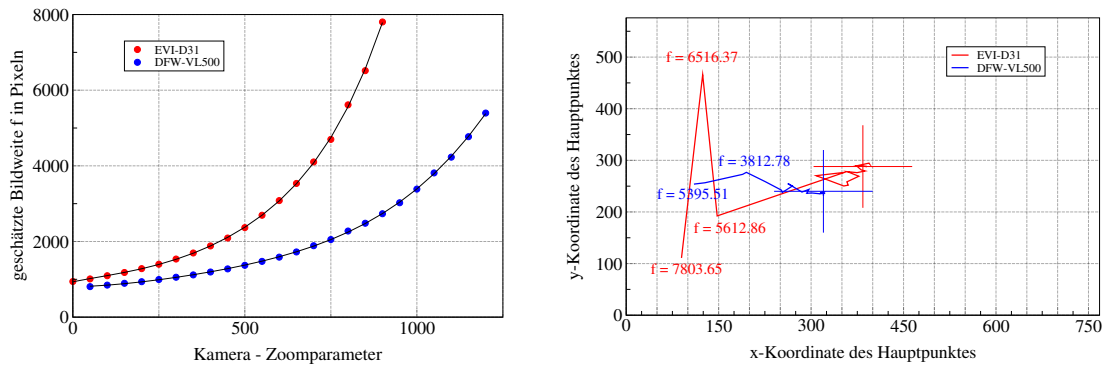


**Abbildung 2.6:** Offline-Kalibrierung: verwendetes Kalibriermuster (links) und Ergebnisse der Farbsegmentierung (mittig) und Eckpunkt-Detektion für die Quadrate des Musters (rechts).

Ebenen. Anschließend wird zu jeder Farbregion diejenige Harris-Ecke gesucht, die das Kalibriermuster der jeweiligen Ebene dort begrenzt und damit das Raster definiert. Dies geschieht, indem zunächst die Schwerpunkte aller Farbregionen berechnet und die jeweils vier nächstgelegenen Harris-Ecken als potenzielle Kandidaten ausgewählt werden. Durch eine senkrechte Projektion dieser vier Punkte auf eine Hilfslinie, die aus einer direkten Verbindung benachbarter Regionenschwerpunkte resultiert (in Abb. 2.6(c) gelb dargestellt), wird schließlich die gesuchte Ecke identifiziert. Sie liegt auf der vom Mittelpunkt der Ebene abgewandten Seite der Linie und ihre Projektion besitzt den größten Abstand zu deren Mittelpunkt. Auf Basis des durch diese Punkte definierten Rasters kann dann zu jedem 3D-Punkt des Kalibrieramusters die korrespondierende 2D-Harris-Ecke im Bild gesucht werden. Ein vorgegebener Maximalabstand zu den durch das Raster implizierten Eckenpositionen soll dabei das Risiko fehlerhafter Zuordnungen vermindern.

Bei der Offline-Kalibrierung hat sich zunächst gezeigt, dass bei kleinen bis mittleren Zoomparametern der Kameras im Regelfall alle 256 Eckpunkte vom Harris-Corner-Detektor robust detektiert werden. Mit steigenden Werten für die Parameter, die mit großen Bildweiten korrelieren, hat sich die Kalibrierung jedoch als zunehmend schwieriger herausgestellt. Insbesondere eine abnehmende Bildqualität ist dabei als Ursache für schlechtere Resultate der Eckendetektion anzuführen. Als Konsequenz wurden nicht die vollständigen Parameterbereiche beider Kameras kalibriert, sondern nur Teilintervalle zwischen 0 und 900 bei der „EVI-D31“- und zwischen 50 und 1200 bei den „DFW-VL500“-Kameras. Die ausgeklammerten Bereiche, die großen Zoomstufen entsprechen, sind jedoch im Kontext dieser Arbeit nur von geringer Bedeutung, da die Kameras vornehmlich im Inneren von Gebäuden eingesetzt werden.

Die Detektion der Harris-Ecken wurde grundsätzlich, abgesehen von einer abnehmenden Bildqualität mit steigender Bildweite, auch durch Linsenverzerrungseffekte negativ beeinflusst. So wichen die 2D-Positionen der Ecken mitunter deutlich von den durch das Raster prädierten Positionen ab, was teilweise nur über eine empirisch gewählte, bildspezifische und auch lokal variierende Reskalierung der Punkteraster auszugleichen war. Abbildung 2.6(c) zeigt ein exemplarisches Detektionsergebnis. Die blauen Kreuze markieren die durch das Raster implizierten Positionen der einzelnen Ecken der Quadrate, während die roten Linien auf die jeweils zugeordneten Harris-Ecken verweisen.



**Abbildung 2.7:** Offline-Kalibrierungsergebnisse für beide Kamertypen: Links sind die Abbildungsfunktionen der Zoomparameter auf die korrespondierenden Bildweiten gezeigt, während rechts die Ergebnisse der Hauptpunktschätzungen zu sehen sind. Dabei fallen insbesondere die großen Abweichungen der Werte von den Bildzentren mit einer zunehmenden Bildweite  $f$  auf.

Die ermittelten Werte für  $\alpha_x$  und  $\alpha_y$  weisen für beide Kamertypen einen linearen Korrelationskoeffizienten von nahezu 1 auf, und interpolierte Regressionsgeraden besitzen in beiden Fällen eine Steigung von 1,0032 bzw. 1,0023. Dies rechtfertigt die Annahme quadratischer CCD-Elemente für beide Kamertypen, so dass jeweils eine einzelne Funktion genügt, um den Zusammenhang zwischen den Zoomparametern und der Bildweite zu beschreiben. Letztere wird als Mittelwert von  $\alpha_x$  und  $\alpha_y$  definiert. Aufgetragen gegen die Hardwareparameter der Kameras (Abb. 2.7, links) implizieren die Werte deutlich einen polynomialen Abbildungszusammenhang, wobei für den Kamertyp „EVI-D31“ ein interpoliertes Polynom fünften Grades den geringsten mittleren quadratischen Approximationsfehler aufweist, während die Werte für die Kameras des Typs „DFW-VL500“ durch ein Polynom vierten Grades am besten beschrieben werden können (schwarze Linien in Abb. 2.7, links). Die rechts in Abbildung 2.7 dargestellten Schätzungen der Hauptpunkte beider Kameras zeigen jeweils für kleine und mittlere Hardwareparameter eine deutliche Übereinstimmung mit den durch die Fadenkreuze markierten Bildzentren (der Kalibrierung der „EVI-D31“-Kamera lagen Bilder im Format  $768 \times 576$  zu Grunde, mit den „DFW-VL500“-Kameras wurden Bilder im Format  $640 \times 480$  aufgenommen). Je größer die Hardwareparameter jedoch gewählt werden, desto größer werden die Abweichungen. Gleichzeitig nimmt die numerische Stabilität innerhalb der ermittelten Werte ab.

Vergleichbare Effekte sind auch bei Untersuchungen zur Reproduzierbarkeit der Kalibrierungsergebnisse zu beobachten, in deren Rahmen für ausgewählte Kameraeinstellungen wiederholt Kalibrierschritte durchgeführt wurden. Bei diesen Analysen zeigt sich mit zunehmenden Bildweiten  $f$  bei beiden Kameras eine leicht ansteigende Varianz in den Ergebnissen, wobei die Standardabweichung im Verhältnis zu den geschätzten mittleren Werten jedoch zumeist noch immer deutlich unter 10% liegt. Lediglich die Hauptpunktkoordinaten fallen für große  $f$  aus diesem Rahmen (vgl. auch Abb. 2.7, rechts).

Die aus dieser Offline-Kalibrierung resultierenden, intrinsischen Parameter der Kameras, d.h. insbesondere auch die Bildweite, können Online-Algorithmen grundsätzlich in Form von Look-Up-Tabellen oder direkt über die interpolierten Funktionen zugänglich gemacht werden. Dies bindet die Algorithmen jedoch an eine Verwendung von Kamertypen, die ein Auslesen ihrer aktuellen Parametereinstellungen zulassen und entsprechend

im Vorfeld einer Offline-Kalibrierung unterzogen werden müssen. Als flexiblere Alternative wurden daher auch die vorgestellten Online-Verfahren erprobt, wobei nur die Kamera des Typs „EVI-D31“ aufgrund der höheren Anzahl rotatorischer Freiheitsgrade zum Einsatz kam. Die direkte Berechnung der Bildweite  $f$  mit dem von Shum und Szeliski vorgeschlagenen Ansatz erscheint dabei grundsätzlich zu instabil. Bei der Online-Kalibrierung mit Hilfe des absolute Conic ist dagegen eine variierende Güte und Stabilität der Ergebnisse in Abhängigkeit von den zu Grunde gelegten Annahmen und Kamerabewegungen zu beobachten. In allen Schätzungen wurden jeweils quadratische CCD-Elemente (begründet durch die Daten der Offline-Kalibrierung) und ein vernachlässigbar kleiner Skew angenommen.

Eine gleichzeitige Schätzung von Hauptpunkt und Bildweite hat sich zunächst als sehr unzuverlässig erwiesen. Dies bestätigt die Ausführungen von de Agapito und Kollegen, wonach eine Verschiebung des Hauptpunktes im Bild oftmals kaum von einer Rotation zu unterscheiden ist [dA00]. Die vorherige, und aufgrund der Offline-Kalibrierung auch gerechtfertigte Fixierung des Hauptpunktes im Bildzentrum führt demgegenüber zwar zu einer deutlichen Stabilisierung der Resultate, die Bildweite wird aber auch mit dieser Voraussetzung im Allgemeinen deutlich überschätzt. Die Werte weisen zudem eine hohe Varianz auf. Dies kann darauf hindeuten, dass die durchgeführten Kamerabewegungen möglicherweise keine hinreichende Grundlage für eine robuste Kalibrierung darstellen und daher Singularitäten und Inkonsistenzen in den verwendeten Homographien vorliegen. De Agapito stellt in dieser Hinsicht insbesondere Rotationen um einzelne Achsen und nur kleine Rotationswinkel als problematisch heraus, da in diesen Fällen eine Bildweitenänderung oftmals nicht von einer Rotation unterschieden werden kann.

Eine Verbesserung der Qualität der aus einer Autokalibrierung resultierenden Kameraparameter ist einerseits bei größeren Kamerarotationen zu erwarten. Dadurch nimmt jedoch auch die Komplexität der Homographieschätzungen zu (Kap. 3), was wiederum eine Erhöhung des dortigen Fehlerrisikos und folglich eine weniger zuverlässige Datengrundlage nach sich zieht. Andererseits wurden bislang Effekte einer Linsenverzerrung in keiner Weise berücksichtigt, so dass auch dadurch verbesserte Resultate zu erzielen sein dürften [Tor00a]. Die Güte einer Autokalibrierung mit den diskutierten Ansätzen hängt allerdings in jedem Fall eng mit der Qualität der geschätzten Homographien zusammen. Da Schätzfehler insbesondere bei einer Online-Bildregistrierung, wie sie im Rahmen dieser Arbeit angestrebt wird, niemals vollständig ausgeschlossen werden können, sind Fehler in einer Bildweitenbestimmung mit Autokalibrierverfahren kaum zu vermeiden. In der Mosaikbildberechnung wird daher einstweilen von einer Bestimmung der Bildweite durch eine Autokalibrierung abgesehen und auf die insgesamt stabileren und auch vermeintlich zuverlässigeren Daten aus der Offline-Kalibrierung zurückgegriffen. Die für die Multi-Mosaikbilder notwendigen Bildweiten werden somit aus den jeweils aktuellen Hardwareparametern und den Interpolationspolynomen berechnet, die im Rahmen der Offline-Kalibrierung für die Abbildung der Zoomparameter auf die korrespondierenden Bildweiten bestimmt wurden. Auch dabei bleibt allerdings abschließend anzumerken, dass mit den geschätzten Werten zwar gute Ergebnisse erzielt werden konnten, eine explizite Korrektur von Linsenverzerrungen jedoch weitere Verbesserungen erwarten lässt.