Rekonstruktion der Gestalt und Bewegung im Rahmen der Geometrischen Algebra

Diplomarbeit

am Lehrstuhl für Kognitive Systeme
des Instituts für Informatik und Praktische Mathematik
der Christian-Albrechts-Universität zu Kiel
Prof. Dr. Gerald Sommer

vorgelegt von

Vladimir Banarer

Betreuer: Dr. Eduardo Bayro-Corrochano

Mai 1999

Inhaltsverzeichnis

Abbildungsverzeichnis

Einleitung

Im Bereich Computer Vision beschäftigt man sich mit der Verarbeitung von visuellen Daten. Die Art der dabei zu lösenden Probleme ist durch die Spezifik dieser Daten festgelegt – man versucht die visuelle Wahrnehmungseigenschaft des Menschen mit Hilfe des Computers nachzuempfinden. Die Palette an Aufgaben reicht dabei von Mustererkennung bis zu 3-D Rekonstruktion von Objekten.

Eine der wichtigsten Aufgaben der Computer Vision besteht darin, anhand von visuellen Daten mit möglichst wenig zusätzlichen Informationen die räumliche Struktur von Objekten sowie die Position des Beobachters relativ zu diesen Objekten zu finden. Diese Probleme werden Rekonstruktionsprobleme genannt. Anhand von verfügbaren Informationen wird zwischen kalibrierter und unkalibrierter Rekonstruktion unterschieden. Bei der kalibrierten Rekonstruktion braucht man nur die 3-D Repräsentation des Objektes zu finden, da durch die Kalibrierung die Kamerapositionen (Beobachterposition) bekannt sind. Bei der unkalibrierten Rekonstruktion wird dagegen die Information auch über die Kamerapositionen gesucht – man spricht von Rekonstruktion der Gestalt und der Bewegung.

Die verschiedenen Aspekte der unkalibrierten Rekonstruktion wurden von mehreren Autoren behandelt – es existieren verschiedenste Ansätze, die unter der Anwendung von unterschiedlichen algebraischen Modellen benutzt werden. Dazu zählen unter anderem projektive Tiefe [1, 2], projektive Invarianten [3] und Faktorisierungsmethoden [4, 5, 6]. Bei diesen Methoden werden die Berechnungen unter dem Einsatz von Matrizen- bzw. Grassmann-Caley Algebra durchgeführt.

Jedes der obengenannten Verfahren bringt Vor- und Nachteile mit sich. So beschränken sich die Faktorisierungsmethoden nur auf die Rekonstruktion von sichtbaren Punkten, sind aber relativ stabil gegenüber dem Einfluß von Rauschen. Die invariantenbasierten Verfahren berücksichtigen oft die Szenengeometrie und ermöglichen die Rekonstruktion von verdeckten Punkten, wenn das a priori Wissen über die Szene vorhanden ist. Allerdings sind sie ziemlich rauschempfindlich. Fast alle Methoden sind aber nicht ausreichend analysiert worden: die Autoren führen die Experimente mit künstlichen Daten durch – die Ergebnisse sind sehr selten mit echten Daten getestet.

Um das Problem der unkalibrierten Rekonstruktion vielseitig behandeln zu können, sollte man die obengenannten Methoden kombinieren und vervollständigen. Dazu ist es notwendig, eine Sprache zu finden bzw. zu entwickeln, in deren Rahmen die gesamten Konzepte einheitlich und konsistent ausgedrückt werden können. Einen passenden mathematischen Zugang für die Lösung solcher Probleme bietet die Geometrische Algebra [7, 8].

Die Geometrische Algebra ist eine Erweiterung der Grassmann- [9] und der Clifford- [10] Algebra. Mit ihrer Hilfe kann man koordinatenunabhängig Geometrie beschreiben. Weiterhin besitzt die Geometrische Algebra solche Operationen wie *join* und *meet*, die es ermöglichen, die geometrischen Beziehungen zwischen den Entitäten verschiedenen Grades einfach aufzustellen. Aufgrund dieser Eigenschaften könnte die Geometrische Algebra sich zum universellen Werkzeug zur Beschreibung und Lösung von verschiedenen Problemen in Computer Vision entwickeln.

Das Ziel dieser Arbeit ist es, die bekannten Methoden für die unkalibrierte Rekonstruktion in der Sprache der Geometrischen Algebra umzuformulieren und zu kombinieren und auf dieser Basis neue Zugänge für die Behandlung dieses Problems zu entwickeln. Dabei sollen die Vorteile der Geometrischen Algebra für die Beschreibung von geometrischen Zusammenhängen der Entitäten im Raum und in der Bildebene demonstriert werden.

Gliederung der Arbeit

Am Anfang steht die Einführung in die Geometrische Algebra. Diesem Thema widmet sich Kapitel 1. Es wird dem Leser ein Einblick in die Grundlagen dieser Algebra verschafft. Weiterhin werden die wichtigsten Operationen zur Manipulation von Algebraentitäten – join und meet – vorgestellt. Danach wird erläutert, wie man mit diesen Operationen neue Objekte konstruieren kann.

In Kapitel 2 wird die Kamerageometrie in der Sprache der Geometrischen Algebra beschrieben. Zunächst wird das Kameramodell untersucht. Danach wird die Stereo- und Trifokalgeometrie behandelt. Dabei werden die Begriffe der Fundamentalmatrix und des Trifokaltensors eingeführt.

Kapitel 3 widmet sich der Invarianztheorie. Es wird gezeigt, wie man mit a priori Wissen über die Punktekonstellation neue Invarianten im Raum aus der allgemeinen Formel ableiten kann. Weiterhin werden die Wege zur Invariantenberechnung mit Bildaten von mehreren Kameras präsentiert. Die Gleichungen von Korrekturkoeffizienten zur Berechnung von Invarianten mit echten Daten werden aufgestellt.

Den theoretischen Kern dieser Arbeit bildet Kapitel 4. Die zuvor erwähnten Methoden zur unkalibrierten Rekonstruktion werden mittels Geometrischer Algebra aufgeschrieben. Die SVD-Rekonstruktion wird um das Verfahren zu der Abschätzung von projektiven Tiefen mit Invarianten erweitert. Weiterhin wird ein Alogrithmus aufgestellt, um bei der Invariantenrekonstruktion die optischen Zentren von Kameras zu finden (Rekonstruktion der Bewegung). Außerdem wird gezeigt, wie man anhand der Bildanalyse oder des a priori Wissens die Rekonstruktion um die Punkte erweitern kann, die im Bild nicht sichtbar sind. Anschließend werden Wege zur Transformation der projektiv rekonstruierten Szene in die euklidische Basis vorgestellt.

Weitere Einsatzmöglichkeiten der Geometrischen Algebra und der Invarianztheorie werden in Kapitel 5 diskutiert. Das Agieren mit der Umgebung, ohne diese komplett zu rekonstruieren, wird anhand eines Beispiels illustriert. Dabei handelt es sich um eine Greifaufgabe, die an einen Roboter gestellt wird.

Kapitel 6 ist ganz experimentellen Untersuchungen der zuvor vorgestellten Methoden gewidmet, wobei zuerst die Experimente mit künstlich erzeugten Daten durchgeführt werden. Es wird das Verhalten von verschiedenen Methoden unter der Einwirkung von Rauschen getestet. Schließlich werden die echten Daten für die Rekonstruktion benutzt.

Kapitel 7 enthält die Zusammenfassung und Diskussion der Resultate dieser Arbeit.

Kapitel 1

Einführung in die Geometrische Algebra

Dieses Kapitel führt in die Begriffe der Geometrischen Algebra ein. Es werden die Möglichkeiten dieser Algebra gezeigt, die eine gute Basis zur Beschreibung von geometrischen Entitäten wie Punkten, Geraden und Ebenen bietet. Weiterhin werden die Methoden zur Beschreibung der Beziehungen zwischen solchen Entitäten mittels der Geometrischen Algebra vorgestellt.

1.1 Grundlegende Definitionen

Man bezeichnet mit $\mathcal{G}_n$ die n-dimensionale Geometrische Algebra – einen geordneten linearen Raum. Dieser Raum hat folgende Eigenschaften:

- Vektoraddition
- Skalarmultiplikation
- eine bis auf Skalarmultiplikation nichtkommutative Vektormultiplikation, die bezüglich Addition assoziativ und distributiv ist, das sogenannte *geometrische* oder *Clifford* Produkt
- jeder Vektor multipliziert mit sich selbst ergibt einen Skalar

Das geometrische Produkt zweier Vektoren $\boldsymbol{a}$ und $\boldsymbol{b}$ kann als die Summe ihrer symmetrischen und antisymmetrischen Komponenten geschrieben werden:

$$\boldsymbol{ab} = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \wedge \boldsymbol{b}, \tag{1.1}$$

wobei das innere Produkt $\boldsymbol{a} \cdot \boldsymbol{b}$ und das äußere Produkt $\boldsymbol{a} \wedge \boldsymbol{b}$ folgendermaßen definiert sind:

$$\boldsymbol{a} \cdot \boldsymbol{b} = \frac{(\boldsymbol{ab} + \boldsymbol{ba})}{2} \tag{1.2}$$

und

$$\boldsymbol{a} \wedge \boldsymbol{b} = \frac{(\boldsymbol{ab} - \boldsymbol{ba})}{2}. \tag{1.3}$$

Das innere Produkt zweier Vektoren ist das Standardskalarprodukt und ergibt einen Skalar.

Das äußere Produkt $\boldsymbol{a} \wedge \boldsymbol{b}$ zweier Vektoren ist eine neue Größe, die *Bivektor* genannt wird. Man stellt sich einen Bivektor als eine gerichtete Fläche in der Ebene vor, die $\boldsymbol{a}$ und $\boldsymbol{b}$ enthält. Das Produkt $\boldsymbol{b} \wedge \boldsymbol{a}$ hat das umgekehrte Vorzeichen (die umgekehrte Richtung), da das äußere Produkt antikommutativ ist, was aus (1.3) folgt.

Das äußere Produkt kann man auf höhere Dimensionen verallgemeinern ($\boldsymbol{a} \wedge \boldsymbol{b} \wedge \boldsymbol{c}$ ist z.B. ein *Trivektor* u.s.w.).

Das äußere Produkt von r Vektoren ist ein $\boldsymbol{r}$-*blade* und hat den Grad $\boldsymbol{r}$.

Ein *Multivektor* ist eine lineare Kombination aus Elementen verschiedenen Grades. Der homogene Anteil vom Grad $\boldsymbol{r}$ eines Multivektors $\boldsymbol{M}$ wird mit $\langle \boldsymbol{M} \rangle_{\boldsymbol{r}}$ bezeichnet und *homogener Multivektor* genannt. Für den $\langle . \rangle_{\boldsymbol{r}}$-Operator gelten folgende Regeln:

- $\langle \langle \boldsymbol{M} \rangle_{\boldsymbol{r}} \rangle_{\boldsymbol{r}} = \langle \boldsymbol{M} \rangle_{\boldsymbol{r}}$

- $\langle \boldsymbol{M_1} + \boldsymbol{M_2} \rangle_{\boldsymbol{r}} = \langle \boldsymbol{M_1} \rangle_{\boldsymbol{r}} + \langle \boldsymbol{M_2} \rangle_{\boldsymbol{r}}$

- $\langle \boldsymbol{\lambda M} \rangle_{\boldsymbol{r}} = \boldsymbol{\lambda} \langle \boldsymbol{M} \rangle_{\boldsymbol{r}}$

Also bilden homogene Multivektoren vom gleichen Grad Unterräume, die bezüglich Addition und Skalarmultiplikation abgeschlossen sind.

Das Element des höchsten Grades in einem Raum heißt der *Pseudoskalar*. Der Einheitspseudoskalar wird mit $\boldsymbol{\mathcal{I}}$ bezeichnet und ist von großer Bedeutung für die Dualität.

Die Definition des geometrischen Produktes zweier Vektoren kann man auf höhere Dimensionen erweitern.

Für das geometrische Produkt eines Vektors $\boldsymbol{a}$ mit einem homogenen Multivektor vom Grade $\boldsymbol{r}$, $\boldsymbol{A_r}$, gilt:

$$\boldsymbol{a}\boldsymbol{A_r} = \boldsymbol{a} \cdot \boldsymbol{A_r} + \boldsymbol{a} \wedge \boldsymbol{A_r}, \tag{1.4}$$

wobei der Grad des inneren Produktes

$$\boldsymbol{a} \cdot \boldsymbol{A_r} = \langle \boldsymbol{a}\boldsymbol{A_r} \rangle_{\boldsymbol{r-1}} = \frac{1}{2}(\boldsymbol{a}\boldsymbol{A_r} - (-1)^r \boldsymbol{A_r}\boldsymbol{a}) \tag{1.5}$$

um eins kleiner als $\boldsymbol{A_r}$, und der Grad des äußeren Produktes

$$\boldsymbol{a} \wedge \boldsymbol{A_r} = \langle \boldsymbol{a}\boldsymbol{A_r} \rangle_{\boldsymbol{r+1}} = \frac{1}{2}(\boldsymbol{a}\boldsymbol{A_r} + (-1)^r \boldsymbol{A_r}\boldsymbol{a}) \tag{1.6}$$

um eins höher als $\boldsymbol{A_r}$ ist.

Dieses Prinzip kann man auf zwei homogene Multivektoren $\boldsymbol{A_r}$ und $\boldsymbol{B_s}$ verallgemeinern:

$$\boldsymbol{A_r}\boldsymbol{B_s} = \langle \boldsymbol{AB} \rangle_{\boldsymbol{r+s}} + \langle \boldsymbol{AB} \rangle_{\boldsymbol{r+s-2}} + \boldsymbol{...} + \langle \boldsymbol{AB} \rangle_{\boldsymbol{|r-s|}}. \tag{1.7}$$

Die Terme des größten und kleinsten Grades repräsentieren das äußere

$$\boldsymbol{A_r} \wedge \boldsymbol{B_s} = \langle \boldsymbol{AB} \rangle_{\boldsymbol{r+s}}, \tag{1.8}$$

bzw. innere

$$\boldsymbol{A_r} \cdot \boldsymbol{B_s} = \langle \boldsymbol{AB} \rangle_{\boldsymbol{|r-s|}}. \tag{1.9}$$

Produkt.

1.2 Geometrische Algebra eines n-dimensionalen Raumes

Man kann in einem n-dimensionalen Raum eine orthonormale Basis aus Vektoren $\{\sigma_i\}$ $i = 1, ..., n$ mit $\sigma_i \cdot \sigma_j = \delta_{ij}$ definieren. Die Basis der entsprechenden Algebra $\mathcal{G}_n$ hat $\boldsymbol{2^n}$ Elemente:[1]

$$1, \{\sigma_i\}, \{\sigma_i \wedge \sigma_j\}, \{\sigma_i \wedge \sigma_j \wedge \sigma_k\}, ..., \sigma_1 \wedge \sigma_2 \wedge ... \wedge \sigma_n \equiv \boldsymbol{\mathcal{I}}. \tag{1.10}$$

[1] Für je zwei Vektoren einer orthonormalen Basis gilt $\sigma_i \wedge \sigma_j = \sigma_i\sigma_j$. So wird statt $\sigma_i \wedge \sigma_j$ im weiteren $\sigma_i\sigma_j$ benutzt.

Jeder Multivektor kann in den Termen dieser Basis repräsentiert werden.

Eine geometrische Algebra des n-dimensionalen Raumes $\mathcal{G}_{p,q,r}$ wird mit drei Zahlen $\boldsymbol{p, q}$ und $\boldsymbol{r}$ charakterisiert, die für die Anzahl von Basisvektoren stehen, deren Quadrat 1,-1 und 0 ist, und für die $\boldsymbol{p+q+r=n}$ gilt. Die gerade Unteralgebra einer solchen Algebra wird mit $\mathcal{G}^+_{p,q,r}$ bezeichnet.

Das höchstgradige blade-Element der Algebra (1.10) heißt der *Einheitspseudoskalar* und ist von spezieller Bedeutung. Es wird mit $\boldsymbol{\mathcal{I}}$ (oder $\boldsymbol{\imath}$ bis zu der dritten Dimension) bezeichnet. Der Einheitspseudoskalar definiert die Dualitätsoperationen der Algebra, so ist zum Beispiel das Ergebnis der Multiplikation eines Multivektors vom Grade $\boldsymbol{r}$ mit dem Einheitspseudoskalar ein Multivektor des $\boldsymbol{(n-r)}$-ten Grades.

In Computer Vision werden die Beziehungen zwischen den Objekten im 3-D Raum und ihren 2-D Abbildungen untersucht. Man braucht homogene Koordinaten, um die projektiven Transformationen zwischen dem Raum und der Bildebene linear zu beschreiben. So bettet man die 2-D Bildebene in den projektiven Raum $\mathbb{P}^2(\equiv\mathbb{R}^3)$ und den 3-D Raum in $\mathbb{P}^3(\equiv\mathbb{R}^4)$ ein. Aus diesem Grund wollen wir die Geometrischen Algebren $\mathcal{G}_3$ und $\mathcal{G}_4$ der 2-D und 3-D projektiven Räume näher untersuchen.

1.2.1 Geometrische Algebra des 3-D Raumes

Die Basis für die Geometrische Algebra $\mathcal{G}_{3,0,0}$ des 3-D Raumes hat $2^3 = 8$ Elemente:

- 1 – Skalar,
- $\sigma_1, \sigma_2, \sigma_3$ – 3 Vektoren,
- $\sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1$ – 3 Bivektoren,
- $\sigma_1\sigma_2\sigma_3 \equiv \boldsymbol{\imath}$ – Einheitspseudoskalar.

Diese Algebra hat folgende Eigenschaften:

$$\begin{aligned}(\sigma_i\sigma_j)\sigma_i &= -\sigma_j,\\ (\sigma_i\sigma_j)\sigma_j &= \sigma_i,\\ (\sigma_i\sigma_j)\sigma_k &= \sigma_i\sigma_j\sigma_k,\end{aligned} \tag{1.11}$$

und

$$(\sigma_1\sigma_2\sigma_3)\sigma_i = \sigma_i(\sigma_1\sigma_2\sigma_3), \tag{1.12}$$

und

$$(\sigma_1\sigma_2\sigma_3)^2 = \sigma_1\sigma_2\sigma_3\sigma_1\sigma_2\sigma_3 = \sigma_1\sigma_2\sigma_1\sigma_2\sigma_3^2 = -1. \tag{1.13}$$

Die obigen Gleichungen haben folgende Interpretation:

- (1.11) – ein einfacher Bivektor rotiert einen Vektor um 90 Grad in der Ebene des Bivektors, oder bildet ein Volumen falls der Vektor senkrecht zur Ebene ist

- (1.12) – der Trivektor ist kommutativ bezüglich aller Vektoren und somit auch bezüglich aller Multivektoren

1.2.2 4-D Geometrische Algebra für den projektiven Raum

In einer Geometrischen Algebra $\boldsymbol{\mathcal{G}_{p,q,r}}$ identifiziert man $\boldsymbol{p}$ und $\boldsymbol{q}$ mit der Dimension des maximalen Unterraumes mit der positiven bzw. negativen Signatur.[2] Im Fall von $\mathbb{P}^2$ wurde die Algebra $\mathcal{G}_{3,0,0}$ mit der euklidischen Standardsignatur gewählt.[3] Um konsistent zu bleiben, wird für den projektiven Raum $\mathbb{P}^3$ eine 4-dimensionale Geometrische Algebra gewählt, die dieselbe Signatur besitzt. Diese (negative) Signatur besitzt die Geometrische Algebra $\mathcal{G}_{1,3,0}$ mit der Basis

- 1 – Skalar,

- $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ – 4 Vektoren,

- $\gamma_2\gamma_3, \gamma_3\gamma_1, \gamma_1\gamma_2, \gamma_4\gamma_1, \gamma_4\gamma_2, \gamma_4\gamma_3$ – 6 Bivektoren,

- $\boldsymbol{\mathcal{I}}\gamma_1, \boldsymbol{\mathcal{I}}\gamma_2, \boldsymbol{\mathcal{I}}\gamma_3, \boldsymbol{\mathcal{I}}\gamma_4$ – 4 Trivektoren,

- $\boldsymbol{\mathcal{I}}$ – Einheitspseudoskalar,

wobei $\gamma_4^2 = 1, \gamma_k^2 = -1$ für $k = 1, 2, 3$ ist.

Der Pseudoskalar ist $\boldsymbol{\mathcal{I}} = \gamma_1\gamma_2\gamma_3\gamma_4$ mit

$$\boldsymbol{\mathcal{I}}^2 = (\gamma_1\gamma_2\gamma_3\gamma_4)(\gamma_1\gamma_2\gamma_3\gamma_4) = (\gamma_2\gamma_3\gamma_4)(\gamma_2\gamma_3\gamma_4) = -(\gamma_3\gamma_4)(\gamma_3\gamma_4) = -1. \tag{1.14}$$

Der vierte Basisvektor γ_4 wird als die gewählte Richtung interpretiert. Dieses Element hilft, die Multivektoren aus 4-D und 3-D zu assoziieren ([11]):

$$\boldsymbol{x} = \frac{\boldsymbol{X} \wedge \boldsymbol{\gamma_4}}{\boldsymbol{X} \cdot \boldsymbol{\gamma_4}}, \tag{1.15}$$

wobei $\boldsymbol{x}$ ein Vektor aus $\mathbb{P}^2$ und $\boldsymbol{X}$ ein Vektor aus $\mathbb{P}^3$ ist. Eine ausführliche Beschreibung dieses Konzeptes (*projektiver split*) kann z.B. in [13] gefunden werden.

[2] Die Signatur eines Vektors $\boldsymbol{a}$ ist positiv, negativ oder Null, wenn $\boldsymbol{a}^2 >, <, = 0$ ist.

[3] Die Signatur des Raumes wird durch das Vorzeichen des Quadrats vom Einheitspseudoskalar bestimmt. Das Quadrat jedes Vektors (also auch jedes Basisvektors) in einem euklidischen Raum ergibt immer einen positiven Skalar. So wird der n-dimensionale euklidische Raum mit $\mathcal{G}_{n,0,0}$ beschrieben. Für $\mathcal{G}_{3,0,0}$ ist $\boldsymbol{\imath}^2 = -1$ – also ist die Signatur negativ.

1.3 Inzidenzalgebra

Für die Repräsentation des projektiven Raumes $\mathbb{P}^3$ wurde im vorigen Abschnitt die 4-D Geometrische Algebra $\mathcal{G}_{1,3,0}$ gewählt. Im gegebenen Raum kann jeder Pseudoskalar $\boldsymbol{\mathcal{P}}$ als $\boldsymbol{\mathcal{P}} = \alpha\boldsymbol{\mathcal{I}}$ geschrieben werden, wobei α ein Skalar ist. Wenn man mit $\boldsymbol{\mathcal{I}}^{-1}$ das Inverse von $\boldsymbol{\mathcal{I}}$ bezeichnet ($\boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{I}}^{-1} = 1$), dann gilt:

$$\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{I}}^{-1} = \alpha\boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{I}}^{-1} = \alpha \equiv [\boldsymbol{\mathcal{P}}]. \tag{1.16}$$

Auf diese Weise ist das *Bracket* $[\boldsymbol{\mathcal{P}}]$ des Pseudoskalars $\boldsymbol{\mathcal{P}}$ definiert als der Wert der rechten Multiplikation mit dem Inversen. Das Vorzeichen des Brackets ist unabhängig von der Signatur des Raumes.

Wir führen das Dualitätskonzept durch Definition des *Dualen* $\boldsymbol{A}^*$ eines $\boldsymbol{r}$-blade $\boldsymbol{A}$ ein:

$$\boldsymbol{A}^* = \boldsymbol{A}\boldsymbol{\mathcal{I}}^{-1}. \tag{1.17}$$

Aus dieser Definition folgt für einen n-dimensionalen Raum, daß das Duale eines $\boldsymbol{r}$-blades ein $(\boldsymbol{n}-\boldsymbol{r})$-blade ist (für $\mathbb{P}^2$ z.B. ist das die Dualität zwischen Punkten und Geraden).

Wenn $\boldsymbol{A}$ ein $\boldsymbol{r}$-blade und $\boldsymbol{B}$ ein $\boldsymbol{s}$-blade mit $\boldsymbol{r}+\boldsymbol{s}=\boldsymbol{n}$ ist, dann gilt:

$$[\boldsymbol{A} \wedge \boldsymbol{B}] = (\boldsymbol{A} \wedge \boldsymbol{B})\boldsymbol{\mathcal{I}}^{-1} = \boldsymbol{A} \cdot \boldsymbol{B}^*. \tag{1.18}$$

In einer n-dimensionalen Geometrischen Algebra kann man die *join*-Operation $\bigwedge$ von einem $\boldsymbol{r}$-blade $\boldsymbol{A}$ und einem $\boldsymbol{s}$-blade $\boldsymbol{B}$ definieren:

$$\boldsymbol{A} \bigwedge \boldsymbol{B} = \boldsymbol{A} \wedge \boldsymbol{B}, \tag{1.19}$$

wenn $\boldsymbol{A}$ und $\boldsymbol{B}$ linear unabhängig sind. Anderenfalls ist join als der von diesen blades aufgespannte Unterraum definiert. Die join-Operation kann man als 'gemeinsamer Dividend kleinsten Grades' interpretieren.

Gibt es einen gemeinsamen Faktor von $\boldsymbol{A}$ und $\boldsymbol{B}$ (d.h. existiert ein $\boldsymbol{r}$-blade $\boldsymbol{C}$ mit $\boldsymbol{A} = \boldsymbol{A}'\boldsymbol{C}$ und $\boldsymbol{B} = \boldsymbol{B}'\boldsymbol{C}$), so läßt sich der 'Schnitt' oder die *meet*-Operation $\bigvee$ von $\boldsymbol{A}$ und $\boldsymbol{B}$ wie folgt definieren:

$$(\boldsymbol{A} \bigvee \boldsymbol{B})^* = \boldsymbol{A}^* \bigwedge \boldsymbol{B}^*. \tag{1.20}$$

Das Duale von meet ist join von Dualen.

Für die praktische Anwendung ist es sinnvoller, meet auf folgende Weise zu berechnen:[4]

$$\begin{aligned} \boldsymbol{A} \bigvee \boldsymbol{B} &= (\boldsymbol{A}^* \bigwedge \boldsymbol{B}^*)\boldsymbol{\mathcal{I}} = (\boldsymbol{A}^* \bigwedge \boldsymbol{B}^*)(\boldsymbol{\mathcal{I}}^{-1}\boldsymbol{\mathcal{I}})\boldsymbol{\mathcal{I}} = \\ \pm & \ (\boldsymbol{A}^* \bigwedge \boldsymbol{B}^*)\boldsymbol{\mathcal{I}}^{-1} = \pm(\boldsymbol{A}^* \cdot \boldsymbol{B}). \end{aligned} \tag{1.21}$$

Das Vorzeichen ändert sich in Abhängigkeit von dem Vorzeichen des quadrierten Einheitspseudoskalars $\boldsymbol{\mathcal{I}}^2 = \pm 1$.

1.4 Geometrische Algebra im projektiven Raum

Man benutzt die Konzepte des Dualitätsprinzips sowie join und meet, um die Objekte im 3-D Raum und deren Beziehungen zueinander (im Sinne von Unterräumen) zu beschreiben.

Besteht ein Objekt aus mehreren einfacheren Objekten, so wird die join-Operation zum Aufbauen benutzt. Ist ein Objekt ein Teil mehrerer anderer Objekte, so wird die meet-Operation zur Konstruktion verwendet. Im Folgenden werden diese Möglichkeiten näher erläutert.

1.4.1 Modellierung von geometrischen Entitäten im Raum

Betrachte drei nicht kollineare Punkte $\boldsymbol{X_1}$,$\boldsymbol{X_2}$,$\boldsymbol{X_3}$. Die Gerade $\boldsymbol{L_{12}}$, die die Punkte $\boldsymbol{X_1}$ und $\boldsymbol{X_2}$ verbindet, kann durch den folgenden Bivektor repräsentiert werden:

$$\boldsymbol{L_{12}} = \boldsymbol{X_1} \bigwedge \boldsymbol{X_2}. \tag{1.22}$$

Jeder Punkt $\boldsymbol{X}$, der auf der Gerade $\boldsymbol{L_{12}}$ liegt, erfüllt

$$\boldsymbol{X} \bigwedge \boldsymbol{L_{12}} = \boldsymbol{X} \bigwedge \boldsymbol{X_1} \bigwedge \boldsymbol{X_2} = \boldsymbol{0}. \tag{1.23}$$

Dies ist die Gleichung der Gerade $\boldsymbol{L_{12}}$. Allgemein sagt diese Gleichung, daß $\boldsymbol{X}$ zu dem von $\boldsymbol{X_1}$ und $\boldsymbol{X_2}$ aufgespannten Unterraum gehört.

Analog wird die Ebene $\boldsymbol{\Phi_{123}}$, die die Punkte $\boldsymbol{X_1}$,$\boldsymbol{X_2}$ und $\boldsymbol{X_3}$ enthält, durch den folgenden Trivektor repräsentiert:

$$\boldsymbol{\Phi_{123}} = \boldsymbol{X_1} \bigwedge \boldsymbol{X_2} \bigwedge \boldsymbol{X_3}. \tag{1.24}$$

Die zu dieser Ebene koplanare Punkte erfüllen folgene Geleichung:

$$\boldsymbol{X} \bigwedge \boldsymbol{X_1} \bigwedge \boldsymbol{X_2} \bigwedge \boldsymbol{X_3} = \boldsymbol{0}. \tag{1.25}$$

Auf diese Weise wird üblicherweise die Koplanaritätsbedingung nachgeprüft.

[4]folgt unmittelbar aus (1.18)

1.4.2 Grundberechnungen der Inzidenzalgebra

Für zwei blades $\boldsymbol{A_r = a_1 \wedge a_2 \wedge ... \wedge a_r}$ und $\boldsymbol{B_s = b_1 \wedge b_2 \wedge ... \wedge b_s}$, $(\boldsymbol{s \leq r})$ gilt:

$$\boldsymbol{B_s \cdot A_r = \sum_j \varepsilon(j_1 j_2 ... j_r) B_s \cdot (a_{j_1} \wedge a_{j_2} \wedge ... \wedge a_{j_s}) a_{j_{s+1}} \wedge ... \wedge a_{j_r}}, \quad (1.26)$$

wobei die Summe über alle Kombinationen $\boldsymbol{j = (j_1, j_2, ..., j_r)}$ mit $\boldsymbol{j_1 < j_2 < ... < j_s}$ geht. Es gilt $\boldsymbol{\varepsilon(j_1 j_2 ... j_r) = 1}$, falls $\boldsymbol{j}$ eine gerade Permutation ist, und $\boldsymbol{-1}$ für eine ungerade Permutation.

Gerade und Ebene

Betrachte eine Gerade $\boldsymbol{L = X_1 \bigwedge X_2}$ und eine Ebene $\boldsymbol{\Phi = Y_1 \bigwedge Y_2 \bigwedge Y_3}$. Der Schnittpunkt kann durch die Anwendung der meet-Operation ermittelt werden:

$$\boldsymbol{L \bigvee \Phi = (X_1 \bigwedge X_2) \bigvee (Y_1 \bigwedge Y_2 \bigwedge Y_3)}. \quad (1.27)$$

Nach der Definition von meet gilt:

$$\boldsymbol{L \bigvee \Phi = -L^* \cdot \Phi}, \quad (1.28)$$

da in $\mathbb{P}^3$ $(\mathcal{G}_{1,3,0})$ $\boldsymbol{\mathcal{I}}^2 = -1$ gilt. $\boldsymbol{\mathcal{I}}$ ist kommutativ bezüglich Bivektoren sowie antikommutativ bezüglich Vektoren und Trivektoren in $\mathcal{G}_{1,3,0}$ und $\boldsymbol{\mathcal{I}}^{-1} = -\boldsymbol{\mathcal{I}}$. Dies führt zu folgender Gleichung:

$$\boldsymbol{L^* \cdot \Phi = (L\mathcal{I}^{-1}) \cdot \Phi = -(L\mathcal{I}) \cdot \Phi}. \quad (1.29)$$

Nach (1.26) kann man die meet-Operation folgendermaßen aufschreiben:

$$\begin{aligned} &\boldsymbol{L \bigvee \Phi = -L^* \cdot \Phi = (L\mathcal{I}) \cdot (Y_1 \bigwedge Y_2 \bigwedge Y_3) =} \\ &\boldsymbol{\{(L\mathcal{I}) \cdot (Y_2 \bigwedge Y_3)\} Y_1 + \{(L\mathcal{I}) \cdot (Y_3 \bigwedge Y_1)\} Y_2 +} \\ &\boldsymbol{\{(L\mathcal{I}) \cdot (Y_1 \bigwedge Y_2)\} Y_3}. \end{aligned} \quad (1.30)$$

Da $\boldsymbol{(L\mathcal{I}) \cdot (Y_i \bigwedge Y_j)}$ ein Skalar ist, kann man die obige Formel auf folgende Weise weiterentwickeln:

$$\boldsymbol{(L\mathcal{I}) \cdot (Y_2 \bigwedge Y_3) = \langle \mathcal{I}(X_1 \bigwedge X_2)(Y_2 \bigwedge Y_3) \rangle_0 = \mathcal{I}(X_1 \bigwedge X_2 \bigwedge Y_2 \bigwedge Y_3)}.$$

Ist $\boldsymbol{\mathcal{P} = X_1 \bigwedge X_2 \bigwedge Y_2 \bigwedge Y_3}$, so ist $[\boldsymbol{\mathcal{P}}] = \boldsymbol{(X_1 \bigwedge X_2 \bigwedge Y_2 \bigwedge Y_3)\mathcal{I}^{-1}}$. Wenn man mit $[\boldsymbol{A_1 A_2 A_3 A_4}]$[5] die Größe des durch die vier Vektoren gebildeten Pseudoskalar bezeichnet, so gilt:

[5]Das Bracket wir durch den Pseudoskalar $\boldsymbol{A_1 \bigwedge A_2 \bigwedge A_3 \bigwedge A_4}$ gebildet. Im Folgenden wird statt $[\boldsymbol{A_1 \bigwedge A_2 \bigwedge A_3 \bigwedge A_4}]$ die Schreibweise $[\boldsymbol{A_1 A_2 A_3 A_4}]$ benutzt, wenn die Art, auf die der Pseudoskalar konstruiert wurde, verständlich ist.

$$\boldsymbol{L} \bigvee \boldsymbol{\Phi} = [\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{Y}_2\boldsymbol{Y}_3]\boldsymbol{Y}_1 + [\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{Y}_3\boldsymbol{Y}_1]\boldsymbol{Y}_2 + [\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{Y}_1\boldsymbol{Y}_2]\boldsymbol{Y}_3. \tag{1.31}$$

Das ist der Schnittpunkt (Vektor) der Geraden und der Ebene im $\mathbb{P}^3$. Die Brackets in obiger Gleichung kann man als die Koordinaten dieses Punktes in Termen der Vektorbasis $\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3\}$ interpretieren.

Zwei Ebenen

Betrachte zwei Ebenen $\boldsymbol{\Phi}_1 = \boldsymbol{X}_1 \wedge \boldsymbol{X}_2 \wedge \boldsymbol{X}_3$ und $\boldsymbol{\Phi}_2 = \boldsymbol{Y}_1 \wedge \boldsymbol{Y}_2 \wedge \boldsymbol{Y}_3$. Der Schnitt dieser Ebenen ist definiert durch

$$\boldsymbol{\Phi}_1 \bigvee \boldsymbol{\Phi}_2 = (\boldsymbol{X}_1 \bigwedge \boldsymbol{X}_2 \bigwedge \boldsymbol{X}_3) \bigvee (\boldsymbol{Y}_1 \bigwedge \boldsymbol{Y}_2 \bigwedge \boldsymbol{Y}_3). \tag{1.32}$$

Analog zu (1.30) gilt:

$$\begin{aligned} &\boldsymbol{\Phi}_1 \bigvee \boldsymbol{\Phi}_2 = -\boldsymbol{\Phi}_1^* \cdot (\boldsymbol{Y}_1 \bigwedge \boldsymbol{Y}_2 \bigwedge \boldsymbol{Y}_3) = \\ &\{(\boldsymbol{\Phi}_1\boldsymbol{\mathcal{I}}) \cdot \boldsymbol{Y}_1\}(\boldsymbol{Y}_2 \bigwedge \boldsymbol{Y}_3) + \{(\boldsymbol{\Phi}_1\boldsymbol{\mathcal{I}}) \cdot \boldsymbol{Y}_2\}(\boldsymbol{Y}_3 \bigwedge \boldsymbol{Y}_1) + \\ &\{(\boldsymbol{\Phi}_1\boldsymbol{\mathcal{I}}) \cdot \boldsymbol{Y}_3\}(\boldsymbol{Y}_1 \bigwedge \boldsymbol{Y}_2). \end{aligned} \tag{1.33}$$

Mit der Argumentation wie im vorigen Abschnitt erhält man

$$(\boldsymbol{\Phi}_1\boldsymbol{\mathcal{I}}) \cdot \boldsymbol{Y}_i = -[\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_3\boldsymbol{Y}_i].$$

Daraus folgt :

$$\begin{aligned} \boldsymbol{\Phi}_1 \bigvee \boldsymbol{\Phi}_2 &= [\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_3\boldsymbol{Y}_1](\boldsymbol{Y}_2 \bigwedge \boldsymbol{Y}_3) \\ &+ [\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_3\boldsymbol{Y}_2](\boldsymbol{Y}_3 \bigwedge \boldsymbol{Y}_1) + [\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_3\boldsymbol{Y}_3](\boldsymbol{Y}_1 \bigwedge \boldsymbol{Y}_2). \end{aligned} \tag{1.34}$$

Das ist die Schnittgerade (Bivektor) zweier Ebenen im $\mathbb{P}^3$.

Zwei Geraden

Zwei Geraden schneiden sich nur, wenn sie koplanar sind, d.h. ihre Repräsentationen $\boldsymbol{L}_A = \boldsymbol{X}_1 \wedge \boldsymbol{X}_2$ und $\boldsymbol{L}_B = \boldsymbol{Y}_1 \wedge \boldsymbol{Y}_2$ erfüllen die Gleichung

$$\boldsymbol{L}_A \bigwedge \boldsymbol{L}_B = \boldsymbol{0}. \tag{1.35}$$

Das heißt, man braucht nur den zweidimensionalen projektiven Raum $\mathbb{P}^2$, der mit der Hilfe des Raumes $\mathbb{R}^3$ beschrieben wird. Der Schnittpunkt ist definiert durch

$$\begin{aligned} &\boldsymbol{L_A} \bigvee \boldsymbol{L_B} = -\boldsymbol{L_A^*} \cdot \boldsymbol{L_B} = (\boldsymbol{L_A} \boldsymbol{\imath}) \cdot (\boldsymbol{Y_1} \bigwedge \boldsymbol{Y_2}) = \\ &((\boldsymbol{L_A} \boldsymbol{\imath}) \cdot \boldsymbol{Y_1}) \boldsymbol{Y_2} - ((\boldsymbol{L_A} \boldsymbol{\imath}) \cdot \boldsymbol{Y_2}) \boldsymbol{Y_1}, \end{aligned} \tag{1.36}$$

wobei $\boldsymbol{\imath}$ der Pseudoskalar in $\mathbb{P}^2$ ist. Analog zu vorherigen Abschnitten gilt:

$$(\boldsymbol{L_A} \boldsymbol{\imath}) \cdot \boldsymbol{Y_i} = \langle \boldsymbol{X_1} \boldsymbol{X_2} \boldsymbol{\imath} \boldsymbol{Y_i} \rangle_0 = \boldsymbol{\imath} \boldsymbol{X_1} \boldsymbol{X_2} \boldsymbol{Y_i} = -[\boldsymbol{X_1} \boldsymbol{X_2} \boldsymbol{Y_i}]. \tag{1.37}$$

So kann der Schnittpunkt beschrieben werden als

$$\boldsymbol{L_A} \bigvee \boldsymbol{L_B} = [\boldsymbol{X_1} \boldsymbol{X_2} \boldsymbol{Y_1}] \boldsymbol{Y_2} - [\boldsymbol{X_1} \boldsymbol{X_2} \boldsymbol{Y_2}] \boldsymbol{Y_1}, \tag{1.38}$$

wobei das Bracket $[\boldsymbol{X}\boldsymbol{Y}\boldsymbol{Z}]$ in $\mathbb{P}^2$ als $(\boldsymbol{X} \bigwedge \boldsymbol{Y} \bigwedge \boldsymbol{Z}) \boldsymbol{\imath}^{-1}$ zu verstehen ist.

Kapitel 2

Kamerageometrie

Viele Methoden der Computer Vision erfordern die Kenntnis der Kamerageometrie. In dieser Arbeit sind diese Kentnisse für die Abschätzung der Invarianten mit den Bilddaten sowie für die Rekonstruktion notwendig. Für diese Zwecke wollen wir die Kamerageometrie in der Sprache der Geometrischen Algebra beschreiben. Die grundlegende Konzepte dieser Beschreibung haben wir in [12] diskutiert. Weitere Einblicke in dieses Thema wurden der Arbeit von E.Bayro-Corrochano und J.Lasenby [14] entnommen.

Dieses Kapitel ist in drei Abschnitte unterteilt. Im ersten Abschnitt wird das Lochkameramodell vorgestellt. Die Bestandteile dieses Modells – das optische Zentrum der Kamera sowie die optischen Strahlen und Ebenen – werden ausführlich beschrieben. Es werden die Gleichungen aufgestellt, die die Projizierung sowohl von Punkten als auch von Geraden in die Bildebene beschreiben. Weiterhin wird gezeigt, wie das Dualitätsprinzip für die Rekonstruktion von optischen Strahlen und Ebenen angewandt werden kann. Im zweiten und dritten Abschnitt werden die Stereo- und Dreikameramodelle behandelt. Es wird die Epipolar- und Trifokalgeometrie erklärt, die besonders wichtig für die Analyse von Punktkorrespondenzen in mehreren Kameras sind. Die Begriffe der Fundamentalmatrix und des Trifokaltensors werden anschaulich anhand der Geometrie in den Termen der Geometrischen Algebra beschrieben.

2.1 Eine Kamera

Betrachte Abbildung 2.1. Die Bildebene $\boldsymbol{\Phi_A}$ ist durch drei nichtkollineare Punkte (Vektoren) $\boldsymbol{A_1}$, $\boldsymbol{A_2}$ und $\boldsymbol{A_3}$ definiert. Das *optische Zentrum* der Kamera liegt im Punkt $\boldsymbol{A_0}$. Die Vektoren $\{\boldsymbol{A_i}\}$ kann man als eine Basis für die Bildebene $\boldsymbol{\Phi_A} = \boldsymbol{A_1} \bigwedge \boldsymbol{A_2} \bigwedge \boldsymbol{A_3}$ wählen, so daß jeder Punkt $\boldsymbol{A'}$ in $\boldsymbol{\Phi_A}$ als lineare Kombination in Termen dieser Basis geschrieben werden kann:

$$\boldsymbol{A'} = \boldsymbol{\alpha_1 A_1} + \boldsymbol{\alpha_2 A_2} + \boldsymbol{\alpha_3 A_3}. \tag{2.1}$$

Man kann für die Bildebene auch eine Bivektorbasis $\{\boldsymbol{L_i^A}\}$ definieren, die die Geraden in $\boldsymbol{\Phi_A}$ aufspannt:

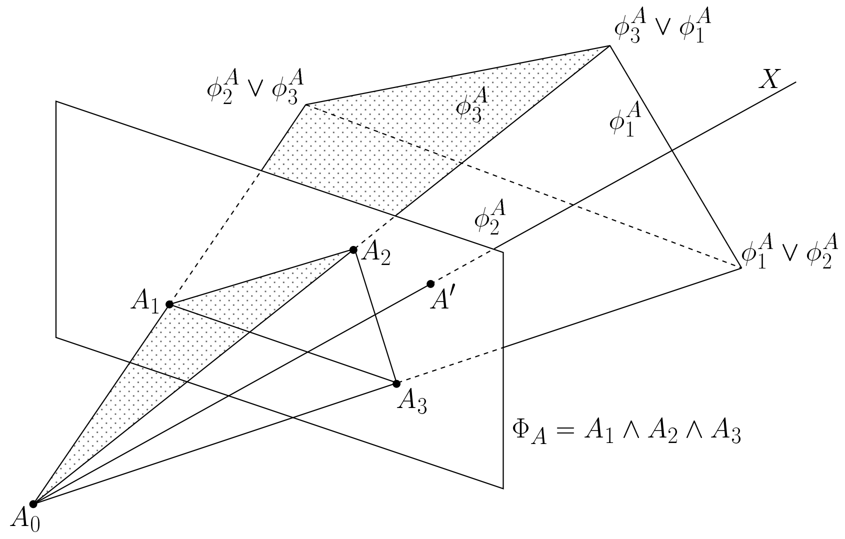


Abbildung 2.1: *Kameramodell für die perspektivische Projektion*

$$\boldsymbol{L_1^A} = \boldsymbol{A_2} \bigwedge \boldsymbol{A_3}, \qquad \boldsymbol{L_2^A} = \boldsymbol{A_3} \bigwedge \boldsymbol{A_1}, \qquad \boldsymbol{L_3^A} = \boldsymbol{A_1} \bigwedge \boldsymbol{A_2}. \tag{2.2}$$

Diese drei Bivektoren zusammen mit dem optischen Zentrum definieren drei Ebenen $\{\boldsymbol{\phi_i^A}\}$, die sogenannten *optischen Ebenen*, wie folgt:

$$\begin{aligned}
\boldsymbol{\phi_1^A} &= \boldsymbol{A_0} \bigwedge \boldsymbol{A_2} \bigwedge \boldsymbol{A_3} = \boldsymbol{A_0} \bigwedge \boldsymbol{L_1^A}, \\
\boldsymbol{\phi_2^A} &= \boldsymbol{A_0} \bigwedge \boldsymbol{A_3} \bigwedge \boldsymbol{A_1} = \boldsymbol{A_0} \bigwedge \boldsymbol{L_2^A}, \\
\boldsymbol{\phi_3^A} &= \boldsymbol{A_0} \bigwedge \boldsymbol{A_1} \bigwedge \boldsymbol{A_2} = \boldsymbol{A_0} \bigwedge \boldsymbol{L_3^A}.
\end{aligned} \tag{2.3}$$

Jede solche Ebene ist ein Trivektor und kann als

$$\boldsymbol{\phi_i^A} = \boldsymbol{t_{i1}(\mathcal{I}\gamma_1) + t_{i2}(\mathcal{I}\gamma_2) + t_{i3}(\mathcal{I}\gamma_3) - t_{i4}(\mathcal{I}\gamma_4)} \tag{2.4}$$

geschrieben werden.[1]

Die optischen Ebenen schneiden die Bildebene entlang der Geraden $\{\boldsymbol{L_i^A}\}$. Weiterhin bilden die Geraden, die als Schnitt der optischen Ebenen entstehen, eine Bivektorbasis für *optische Strahlen* in $\mathbb{R}^4$:

$$\begin{aligned}
\boldsymbol{L_{A1}} &= \boldsymbol{\phi_2^A} \bigvee \boldsymbol{\phi_3^A} = \boldsymbol{A_0} \bigwedge \boldsymbol{A_1}, \\
\boldsymbol{L_{A2}} &= \boldsymbol{\phi_3^A} \bigvee \boldsymbol{\phi_1^A} = \boldsymbol{A_0} \bigwedge \boldsymbol{A_2}, \\
\boldsymbol{L_{A3}} &= \boldsymbol{\phi_1^A} \bigvee \boldsymbol{\phi_2^A} = \boldsymbol{A_0} \bigwedge \boldsymbol{A_3},
\end{aligned} \tag{2.5}$$

[1] Die Vorzeichen von Trivektorbasiselementen für die Darstellung von optischen Ebenen wählen wir in Abhängigkeit von der zu Grunde liegenden Algebra $\mathcal{G}_{p,q,r}$. Es werden die zu dem Vorzeichen des Quadrates von jeweiligen Basiselement Inverse genommen. Im Fall von $\mathcal{G}_{1,3,0}$ sind das folgende Vorzeichen $+\ +\ +\ -$.

so daß jeder optische Strahl, der durch einen Punkt $\boldsymbol{X}$ erzeugt wird, als eine lineare Kombination der Basisstrahlen geschrieben werden kann:

$$\boldsymbol{A}_0 \bigwedge \boldsymbol{X} = \sum_{i=1}^{3} \mathrm{x_i} \boldsymbol{L}_{Ai}.$$

Mit diesen Erkenntnissen kann man die Projektionsmatrix einer Kamera in Termen der Geometrischen Algebra aufschreiben.

Die Projektion $\boldsymbol{x} = (\mathrm{x_1}, \mathrm{x_2}, \mathrm{x_3})^{\mathrm{T}}$ eines Punktes $\boldsymbol{X} = (\mathrm{X_1}, \mathrm{X_2}, \mathrm{X_3}, \mathrm{X_4})^{\mathrm{T}}$ in die Bildebene ist durch den Schnitt des optischen Strahls $\boldsymbol{A_0} \bigwedge \boldsymbol{X}$ mit der Ebene $\boldsymbol{\Phi_A}$ gegeben:

$$\boldsymbol{x} = (\boldsymbol{A}_0 \bigwedge \boldsymbol{X}) \bigvee (\boldsymbol{A}_1 \bigwedge \boldsymbol{A}_2 \bigwedge \boldsymbol{A}_3) = \sum_{i=1}^{4} \mathrm{X_i} \{(\boldsymbol{A}_0 \bigwedge \gamma_i) \bigvee (\boldsymbol{A}_1 \bigwedge \boldsymbol{A}_2 \bigwedge \boldsymbol{A}_3)\}. \tag{2.6}$$

Die Gleichung (2.6) kann man auf folgende Weise umschreiben:

$$\begin{aligned} \boldsymbol{x} &= \sum_{i=1}^{4} \mathrm{X_i} \{[\boldsymbol{A}_0 \bigwedge \gamma_i \bigwedge \boldsymbol{A}_2 \bigwedge \boldsymbol{A}_3] \boldsymbol{A}_1 \\ &+ [\boldsymbol{A}_0 \bigwedge \gamma_i \bigwedge \boldsymbol{A}_3 \bigwedge \boldsymbol{A}_1] \boldsymbol{A}_2 + [\boldsymbol{A}_0 \bigwedge \gamma_j \bigwedge \boldsymbol{A}_1 \bigwedge \boldsymbol{A}_2] \boldsymbol{A}_3\}. \end{aligned} \tag{2.7}$$

Nach (2.1) ist dann die Projektion des Punktes $\boldsymbol{X}$ in Termen der Basis $\{\boldsymbol{A}_i\}$ folgendermaßen definiert:

$$\mathrm{x_i} = \sum_{j=1}^{4} \boldsymbol{P}_{ij} \mathrm{X_j},$$

mit

$$\boldsymbol{P}_{ij} = [\boldsymbol{A}_0 \bigwedge \gamma_j \bigwedge \boldsymbol{L}_i^{\boldsymbol{A}}] = [\boldsymbol{\phi}_i^{\boldsymbol{A}} \bigwedge \gamma_j] = -\boldsymbol{t}_{ij}. \tag{2.8}$$

Die Matrix $\boldsymbol{P}_{3\times 4}$ bildet die Raumpunkte aus dem $\mathbb{P}^3$ in die Bildpunkte aus $\mathbb{P}^2$ ab, und heißt die *Projektionsmatrix*.

Wenn wir die zu den optischen Ebenen $\{\boldsymbol{\phi}_i^{\boldsymbol{A}}\}$ dualen Vektoren $\{\boldsymbol{\phi}_{\boldsymbol{A}}^i\}$ definieren ($\boldsymbol{\phi}_{\boldsymbol{A}}^i = \boldsymbol{\phi}_i^{\boldsymbol{A}} \boldsymbol{\mathcal{I}}^{-1}$), so gilt in $\mathcal{G}_{1,3,0}$:

$$\boldsymbol{\phi}_{\boldsymbol{A}}^i = -\boldsymbol{\phi}_i^{\boldsymbol{A}} \boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{I}} \boldsymbol{\phi}_i^{\boldsymbol{A}} = -(\boldsymbol{t}_{i1}\gamma_1 + \boldsymbol{t}_{i2}\gamma_2 + \boldsymbol{t}_{i3}\gamma_3 - \boldsymbol{t}_{i4}\gamma_4). \tag{2.9}$$

Auf diese Weise kann man den projizierten Punkt $\boldsymbol{x} = \sum_{i=1}^{3} \mathrm{x_i}\boldsymbol{A}_i$ folgendermaßen aufschreiben:

$$\boldsymbol{x} = \sum_{i=1}^{3} (\boldsymbol{X} \cdot \boldsymbol{\phi}_A^i)\boldsymbol{A}_i, \tag{2.10}$$

d.h. die Koeffizienten in der Bildebene sind definiert durch die Projektion des Punktes $\boldsymbol{X}$ in die zu den optischen Ebenen dualen Vektoren. Das ist äquivalent zur üblichen Matrizendarstellung:[2]

$$\boldsymbol{x} = \begin{bmatrix} \mathrm{x_1} \\ \mathrm{x_2} \\ \mathrm{x_3} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_A^1 \\ \boldsymbol{\phi}_A^2 \\ \boldsymbol{\phi}_A^3 \end{bmatrix} \boldsymbol{X} \equiv \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \end{bmatrix} \begin{bmatrix} \mathrm{X_1} \\ \mathrm{X_2} \\ \mathrm{X_3} \\ \mathrm{X_4} \end{bmatrix} = \boldsymbol{P}\boldsymbol{X}. \tag{2.11}$$

Also sind die Elemente der Projektionsmatrix die Koeffizienten der Dualen der optischen Ebenen.

Ähnlich wird auch die Projektionsmatrix für die Geraden gefunden. Betrachte eine Gerade $\boldsymbol{L} = \boldsymbol{X_1} \bigwedge \boldsymbol{X_2}$, die durch die Punkte $\boldsymbol{X_1}$ und $\boldsymbol{X_2}$ aufgespannt ist. Wenn $\boldsymbol{x_1} = (\boldsymbol{A_0} \bigwedge \boldsymbol{X_1}) \bigvee \boldsymbol{\Phi_A}$ und $\boldsymbol{x_2} = (\boldsymbol{A_0} \bigwedge \boldsymbol{X_2}) \bigvee \boldsymbol{\Phi_A}$ die Projektionen dieser Punkte in die Bildebene sind, so sieht die Geradenabbildung folgendermaßen aus:

$$\boldsymbol{l} = \boldsymbol{x_1} \bigwedge \boldsymbol{x_2}.$$

Wir können $\boldsymbol{l}$ in Bivektorbasistermen der Bildebene $\{\boldsymbol{L_{Ai}}\}$ ausdrücken:

$$\boldsymbol{l} = \sum_{i=1}^{3} \mathrm{l_i}\boldsymbol{L}_i^A. \tag{2.12}$$

Nach (2.10) kann man $\boldsymbol{l}$ auf folgende Weise aufschreiben:

$$\boldsymbol{l} = \boldsymbol{x_1} \bigwedge \boldsymbol{x_2} = \sum_{i,j=1}^{3} (\boldsymbol{X_1} \cdot \boldsymbol{\phi}_A^i)(\boldsymbol{X_2} \cdot \boldsymbol{\phi}_A^j)\boldsymbol{A}_i \bigwedge \boldsymbol{A}_j = \sum_{k=1}^{3} \mathrm{l_k}\boldsymbol{L}_k^A, \tag{2.13}$$

wobei die Geradenkoeffizienten $\{\boldsymbol{l_k}\}$ wie folgt sind

$$\begin{aligned} \boldsymbol{l_1} &= (\boldsymbol{X_1} \cdot \boldsymbol{\phi}_A^2)(\boldsymbol{X_2} \cdot \boldsymbol{\phi}_A^3) - (\boldsymbol{X_1} \cdot \boldsymbol{\phi}_A^3)(\boldsymbol{X_2} \cdot \boldsymbol{\phi}_A^2), \\ \boldsymbol{l_2} &= (\boldsymbol{X_1} \cdot \boldsymbol{\phi}_A^3)(\boldsymbol{X_2} \cdot \boldsymbol{\phi}_A^1) - (\boldsymbol{X_1} \cdot \boldsymbol{\phi}_A^1)(\boldsymbol{X_2} \cdot \boldsymbol{\phi}_A^3), \\ \boldsymbol{l_3} &= (\boldsymbol{X_1} \cdot \boldsymbol{\phi}_A^1)(\boldsymbol{X_2} \cdot \boldsymbol{\phi}_A^2) - (\boldsymbol{X_1} \cdot \boldsymbol{\phi}_A^2)(\boldsymbol{X_2} \cdot \boldsymbol{\phi}_A^1). \end{aligned} \tag{2.14}$$

[2]Die Bedeutung der Koeffizienten der Projektionsmatrix ist in Anhang A erläutert.

Mit (1.20) ergibt sich:

$$\boldsymbol{l}_1 = (\boldsymbol{X}_1 \bigwedge \boldsymbol{X}_2) \cdot (\boldsymbol{\phi}_A^2 \bigwedge \boldsymbol{\phi}_A^3) = (\boldsymbol{X}_1 \bigwedge \boldsymbol{X}_2) \cdot (\boldsymbol{\phi}_2^A \bigvee \boldsymbol{\phi}_3^A)^* = \boldsymbol{L} \cdot \boldsymbol{L}_1^{A*}.$$

Allgemein gilt für jeden Geradenkoeffizient $\boldsymbol{l}_i$:

$$\boldsymbol{l}_i = \boldsymbol{L} \cdot \boldsymbol{L}_i^{A*} = \boldsymbol{L} \cdot \boldsymbol{L}_A^i, \tag{2.15}$$

wobei mit $\{\boldsymbol{L}_A^i\}$ die Dualen zu den Geraden $\{\boldsymbol{L}_i^A\}$ bezeichnet werden. Auf diese Weise wird die Projektion $\boldsymbol{l}$ einer Gerade $\boldsymbol{L}$ in die Bildebene in Termen von Dualen der Basisstrahlen ausgedrückt – die Gleichung für die Projektion lautet:

$$\boldsymbol{l} = \sum_{i=1}^{3} (\boldsymbol{L} \cdot \boldsymbol{L}_A^i) \boldsymbol{L}_i^A. \tag{2.16}$$

Diese Gleichung kann man mit der Hilfe von (2.12) auch in der üblichen Matrixform aufschreiben:

$$\boldsymbol{l} = \begin{bmatrix} \mathrm{l}_1 \\ \mathrm{l}_2 \\ \mathrm{l}_3 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} & u_{15} & u_{16} \\ u_{21} & u_{22} & u_{23} & u_{24} & u_{25} & u_{26} \\ u_{31} & u_{32} & u_{33} & u_{34} & u_{35} & u_{36} \end{bmatrix} \begin{bmatrix} \mathrm{L}_1 \\ \mathrm{L}_2 \\ \mathrm{L}_3 \\ \tilde{\mathrm{L}}_1 \\ \tilde{\mathrm{L}}_2 \\ \tilde{\mathrm{L}}_3 \end{bmatrix} = \boldsymbol{P}_L \bar{\boldsymbol{L}}, \tag{2.17}$$

wobei $\bar{\boldsymbol{L}}$ die Gerade mit den sogenannten *Plücker Koordinaten* $(\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \tilde{\mathbf{L}}_1, \tilde{\mathbf{L}}_2, \tilde{\mathbf{L}}_3)$ ist.

Man kann das Dualitätsprinzip auf die Projektionsgleichungen (2.11) und (2.17) für den projizierten Punkt $\boldsymbol{x}$ bzw. die Gerade $\boldsymbol{l}$ anwenden:

$$\boldsymbol{L}_x = \boldsymbol{A}_0 \bigwedge \boldsymbol{x} = \sum_{i=1}^{3} (\boldsymbol{X} \cdot \boldsymbol{\phi}_A^i) \boldsymbol{A}_0 \bigwedge \boldsymbol{A}_i = \sum_{i=1}^{3} (\boldsymbol{X} \cdot \boldsymbol{\phi}_A^i) \boldsymbol{L}_i^A \equiv \boldsymbol{P}_L^T \boldsymbol{x}, \tag{2.18}$$

$$\boldsymbol{\phi}_l = \boldsymbol{A}_0 \bigwedge \boldsymbol{l} = \sum_{i=1}^{3} (\boldsymbol{L} \cdot \boldsymbol{L}_A^i) \boldsymbol{A}_0 \bigwedge \boldsymbol{L}_i^A = \sum_{i=1}^{3} (\boldsymbol{L} \cdot \boldsymbol{L}_A^i) \boldsymbol{\phi}_i^A \equiv \boldsymbol{P}^T \boldsymbol{l}. \tag{2.19}$$

Dabei erhält man die entsprechenden Dualen – die Gerade $\boldsymbol{L}_x$ (den optischen Strahl) bzw. die Ebene $\boldsymbol{\phi}_l$ (die optische Ebene).

2.2 Zwei Kameras

Abbildung 2.2 zeigt, wie ein Punkt $\boldsymbol{X}$ im Raum in die Bildebenen $\boldsymbol{\Phi}_A$ und $\boldsymbol{\Phi}_B$ projiziert wird. Die Bildpunkte heißen $\boldsymbol{A}'$ bzw. $\boldsymbol{B}'$.

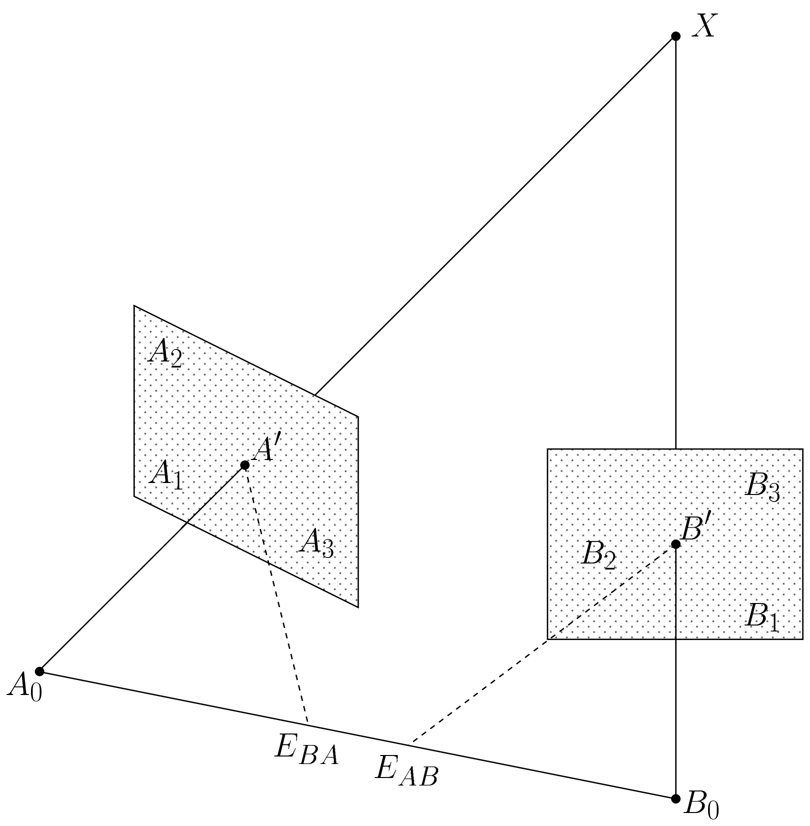


Abbildung 2.2: *Stereogeometrie*

Die sogenannten *Epipole* $\boldsymbol{E_{AB}}$ und $\boldsymbol{E_{BA}}$ sind die Abbildungen des optischen Zentrums einer Kamera in die jeweils andere Kamera.

Die Punkte $\boldsymbol{A_0}$,$\boldsymbol{B_0}$,$\boldsymbol{A'}$,$\boldsymbol{B'}$ sind koplanar – sie liegen in einer Ebene, die *Epipolarebene* genannt wird. Die join-Operation, die auf vier koplanare Punkte in einem dreidimensionalen projektiven Raum angewandt wird, ergibt Null:

$$\boldsymbol{A_0} \bigwedge \boldsymbol{B_0} \bigwedge \boldsymbol{A'} \bigwedge \boldsymbol{B'} = \boldsymbol{0}. \tag{2.20}$$

Wenn man die Bildpunkte $\boldsymbol{A'}$ und $\boldsymbol{B'}$ in Termen der Kamerabasis $\{\boldsymbol{A_i}\}$ bzw. $\{\boldsymbol{B_i}\}$ aufschreibt, so kann man die Gleichung (2.20) wie folgt umformulieren:

$$\sum_{\boldsymbol{i,j=1}}^{\boldsymbol{3}} \boldsymbol{\alpha_i \beta_j} \{\boldsymbol{A_0} \bigwedge \boldsymbol{B_0} \bigwedge \ \boldsymbol{A_i} \bigwedge \ \boldsymbol{B_j}\} = \boldsymbol{0}. \tag{2.21}$$

Wird $\boldsymbol{\tilde{F}_{ij}} = \{\boldsymbol{A_0} \wedge \ \boldsymbol{B_0} \wedge \ \boldsymbol{A_i} \wedge \ \boldsymbol{B_j}\}\boldsymbol{\mathcal{I}^{-1}} = [\boldsymbol{A_0 B_0 A_i B_j}]$ gesetzt, so läßt sich die obige Gleichung auf folgende Weise umschreiben:

$$\sum_{\boldsymbol{i,j=1}}^{\boldsymbol{3}} \boldsymbol{\tilde{F}_{ij} \alpha_i \beta_j} = \boldsymbol{0}. \tag{2.22}$$

Die Gleichung (2.22) beschreibt die bekannte Gleichung für die *Fundamentalmatrix* $\tilde{\boldsymbol{F}}$.

Die $\tilde{\boldsymbol{F}}$-Matrix kann man als eine lineare Funktion interpretieren, die zwei Vektoren auf einen Skalar abbildet:

$$\tilde{\boldsymbol{F}}(\boldsymbol{A}, \boldsymbol{B}) = \{\boldsymbol{A_0} \bigwedge \boldsymbol{B_0} \bigwedge \ \boldsymbol{A} \bigwedge \ \boldsymbol{B}\}\boldsymbol{\mathcal{I}^{-1}} = [\boldsymbol{A_0 B_0 A B}]. \tag{2.23}$$

Auf diese Weise kann man die Koeffizienten der Fundamentalmatrix $\tilde{\boldsymbol{F}}_{\boldsymbol{ij}}$ als eine Funktion von $\tilde{\boldsymbol{F}}(\boldsymbol{A}, \boldsymbol{B})$ definieren:

$$\tilde{\boldsymbol{F}}_{\boldsymbol{ij}} = \tilde{\boldsymbol{F}}(\boldsymbol{A_i}, \boldsymbol{B_j}), \tag{2.24}$$

wobei die $\boldsymbol{A_i}$'s und $\boldsymbol{B_j}$'s die Elemente der jeweiligen Bildebenenbasis sind.

2.3 Drei Kameras

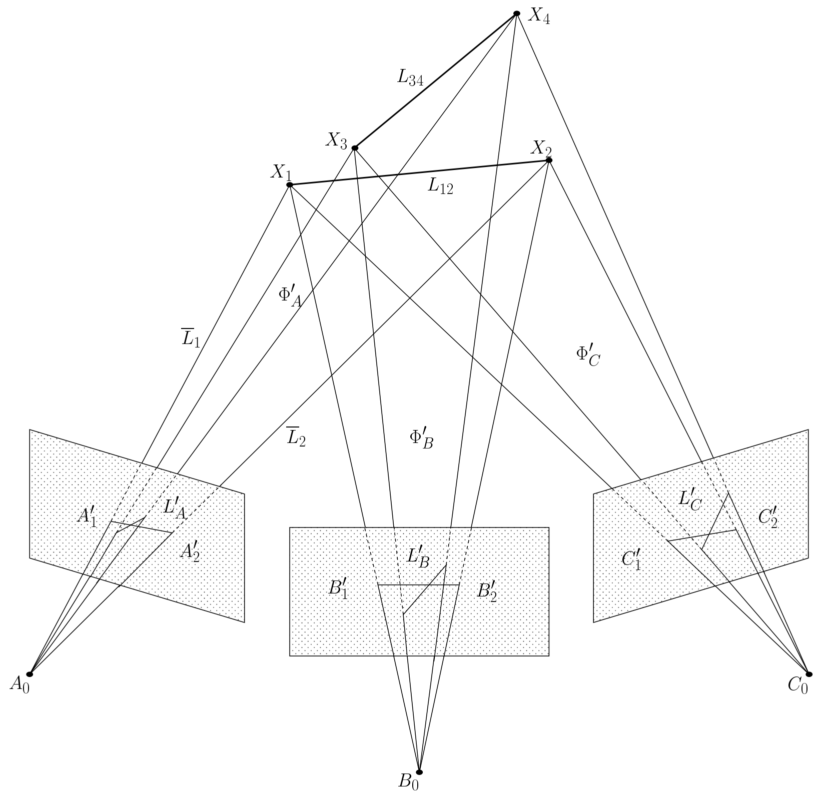


Abbildung 2.3: *Trifokale Geometrie*

Die sogenannte trilineare Zwangsbedingung beschreibt die geometrischen Beziehungen zwischen Punkten und Geraden in drei Kameras. Abbildung 2.3 zeigt drei Bildebenen $\boldsymbol{\Phi_A}$,$\boldsymbol{\Phi_B}$ und $\boldsymbol{\Phi_C}$. Die Punkte $\boldsymbol{X_1}$ und $\boldsymbol{X_2}$ werden in die Kameras $\boldsymbol{A}$,$\boldsymbol{B}$ und $\boldsymbol{C}$ auf die Bildpunkte $\boldsymbol{A'_1}$,$\boldsymbol{A'_2}$,$\boldsymbol{B'_1}$,$\boldsymbol{B'_2}$,$\boldsymbol{C'_1}$,$\boldsymbol{C'_2}$ abgebildet. Die Gerade durch die

Punkte $\boldsymbol{X_1}$ und $\boldsymbol{X_2}$ wird mit $\boldsymbol{L}$ und ihre Abbildungen in die Kameras mit $\boldsymbol{L'_A}$,$\boldsymbol{L'_B}$ und $\boldsymbol{L'_C}$ bezeichnet. Wir definieren drei Ebenen:

$$\boldsymbol{\Phi'_A} = \boldsymbol{A_0} \bigwedge \boldsymbol{A'_1} \bigwedge \boldsymbol{A'_2}, \quad \boldsymbol{\Phi'_B} = \boldsymbol{B_0} \bigwedge \boldsymbol{B'_1} \bigwedge \boldsymbol{B'_2}, \quad \boldsymbol{\Phi'_C} = \boldsymbol{C_0} \bigwedge \boldsymbol{C'_1} \bigwedge \boldsymbol{C'_2}. \tag{2.25}$$

Die Gerade $\boldsymbol{L}$ kann als Schnitt von zwei solcher Ebenen konstruiert werden, z.B.:

$$\boldsymbol{L} = \boldsymbol{\Phi'_B} \bigvee \boldsymbol{\Phi'_C} = (\boldsymbol{B_0} \bigwedge \boldsymbol{L'_B}) \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L'_C}). \tag{2.26}$$

Wenn man $\boldsymbol{L_{A_1}} = \boldsymbol{A_0} \wedge \boldsymbol{A'_1}$ und $\boldsymbol{L_{A_2}} = \boldsymbol{A_0} \wedge \boldsymbol{A'_2}$ setzt, so schneiden sich diese Geraden mit $\boldsymbol{L}$ in den Punkten $\boldsymbol{X_1}$ bzw. $\boldsymbol{X_2}$, d.h., daß die Geraden $\boldsymbol{L_{A_1}}$,$\boldsymbol{L}$ und $\boldsymbol{L_{A_2}}$,$\boldsymbol{L}$ koplanar sind:

$$\boldsymbol{L_{A_1}} \bigwedge \boldsymbol{L} = 0 \qquad \boldsymbol{L_{A_2}} \bigwedge \boldsymbol{L} = 0. \tag{2.27}$$

Diese Gleichung kann man mit (2.26) substituieren:

$$\begin{aligned} (\boldsymbol{A_0} \bigwedge \boldsymbol{A'_1}) \bigwedge \{(\boldsymbol{B_0} \bigwedge \boldsymbol{L'_B}) \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L'_C})\} &= 0, \\ (\boldsymbol{A_0} \bigwedge \boldsymbol{A'_2}) \bigwedge \{(\boldsymbol{B_0} \bigwedge \boldsymbol{L'_B}) \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L'_C})\} &= 0. \end{aligned} \tag{2.28}$$

Analog zur Fundamentalmatrix kann man eine lineare Funktion $\boldsymbol{\tilde{T}}$ definieren, die zwei Geraden und einen Punkt auf einen Skalar abbildet:

$$\boldsymbol{\tilde{T}}(\boldsymbol{A'}, \boldsymbol{L'_B}, \boldsymbol{L'_C}) = [(\boldsymbol{A_0} \bigwedge \boldsymbol{A'}) \bigwedge \{(\boldsymbol{B_0} \bigwedge \boldsymbol{L'_B}) \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L'_C})\}]. \tag{2.29}$$

Gemäß (2.2) definieren wir die Bivektorbasen $\{\boldsymbol{L_i^A}\}$,$\{\boldsymbol{L_i^B}\}$ und $\{\boldsymbol{L_i^C}\}$ für die Kameras $\boldsymbol{A}$,$\boldsymbol{B}$ und $\boldsymbol{C}$ entsprechend. In dieser Basis werden der Punkt $\boldsymbol{A'}$ und die Geraden $\boldsymbol{L'_B}$,$\boldsymbol{L'_C}$ folgendermaßen ausgedrückt:

$$\boldsymbol{A'} = \sum_{i=1}^{3} \boldsymbol{\alpha_i A_i}, \quad \boldsymbol{L'_B} = \sum_{j=1}^{3} \mathrm{l}_\mathrm{j}^\mathrm{B} \boldsymbol{L_j^B}, \quad \boldsymbol{L'_C} = \sum_{k=1}^{3} \mathrm{l}_\mathrm{k}^\mathrm{C} \boldsymbol{L_k^C}. \tag{2.30}$$

Wir definieren $\boldsymbol{\tilde{T}_{ijk}}$ als

$$\boldsymbol{\tilde{T}_{ijk}} = \boldsymbol{\tilde{T}}(\boldsymbol{A_i}, \boldsymbol{L_j^B}, \boldsymbol{L_k^C}).$$

Stammen die Bildpunkte und Geraden in allen Kameras von denselben Punkten und Geraden im Raum ab, so gilt nach (2.28):

$$\sum_{i,j,k=1}^{3} \boldsymbol{\tilde{T}_{ijk}} \boldsymbol{\alpha_i} \mathrm{l}_\mathrm{j}^\mathrm{B} \mathrm{l}_\mathrm{k}^\mathrm{C} = 0. \tag{2.31}$$

Die Gleichung (2.31) wird die *trilineare Zwangsbedingung* genannt. Die $\tilde{\boldsymbol{T}}_{\boldsymbol{ijk}}$-Elemente bilden dabei die Koeffizienten des sogenannten $3 \times 3 \times 3$ *Trifokaltensors*.

Mit der Hilfe des Trifokaltensors kann man auch die Beziehungen zwischen den Geraden in drei Kameras untersuchen. Die Gerade $\boldsymbol{L'_A}$ in der Kamera $\boldsymbol{A}$ kann man als

$$\boldsymbol{L'_A} = \boldsymbol{A'_1} \bigwedge \boldsymbol{A'_2} = (\boldsymbol{A_0} \bigwedge \boldsymbol{L}) \bigvee \boldsymbol{\Phi_A} \tag{2.32}$$

aufschreiben.

Mit (2.26) ergibt sich folgende Beziehung für $\boldsymbol{L'_A}$,$\boldsymbol{L'_B}$,$\boldsymbol{L'_C}$:

$$\boldsymbol{L'_A} = (\boldsymbol{A_0} \bigwedge \{(\boldsymbol{B_0} \bigwedge \boldsymbol{L'_B}) \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L'_C})\}) \bigvee \boldsymbol{\Phi_A}. \tag{2.33}$$

Mit Hilfe von (2.30) erhält man

$$\mathrm{l_i^A} \boldsymbol{L_i^A} = \sum_{\boldsymbol{j,k=1}}^{\boldsymbol{3}} [(\boldsymbol{A_0} \bigwedge \boldsymbol{A_i}) \bigwedge \mathrm{l_j^B l_k^C} \{(\boldsymbol{B_0} \bigwedge \boldsymbol{L_j^B}) \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L_k^C})\}] \boldsymbol{L_i^A}. \tag{2.34}$$

Auf diese Weise wird die Beziehung zwischen den Geraden in der Bildebene in Termen des Trifokaltensors bestimmt:

$$\mathrm{l_i^A} = \sum_{\boldsymbol{j,k=1}}^{\boldsymbol{3}} \tilde{\boldsymbol{T}}_{\boldsymbol{ijk}} \mathrm{l_j^B l_k^C}. \tag{2.35}$$

Ist also die Trifokalgeometrie bekannt, so kann man anhand der Abbildungen einer Gerade in zwei Kameras das Bild dieser Gerade auch in der dritten Kamera finden.

Kapitel 3

Invarianten

Unter einer *Invarianten* versteht man eine geometrische Beziehung zwischen den Punkten im Raum, die unter Einwirkung von beliebigen projektiven Transformationen konstant bleibt. Diese Eigenschaft kann für die Bildanalyse sowie für die Rekonstruktion von großem Nutzen sein. In dieser Arbeit werden die Invarianten für die Abschätzung von projektiven Tiefen und für die Rekonstruktion eingesetzt. Deshalb wollen wir ihre Eigenschaften und Berechnungsarten in diesem Kapitel genauer erläutern.

Zunächst wird die allgemeine Gleichung gebildet, die die Beziehung zwischen den Punkten im Raum und in der Bildebene beschreibt. Diese Gleichung wird für die Generierung von Invarianten im Raum benutzt, wenn einige der Punkte eine spezielle Beziehung zueinander haben – z.B. kollinear oder koplanar sind. Einige Beispiele für auf diese Weise erzeugten Invarianten werden gegeben. Am Anfang werden alle Berechnungen im 3-D Raum durchgeführt. Danach werden die Wege zur Invariantenberechnung anhand der Bilddaten beschrieben. Wir zeigen die Möglichkeiten der Invariantenabschätzung mit Punktkorrespondenzen in mehreren Kameras. Dabei werden die Kentnisse über die Kamerageometrie aus dem vorhergegangenen Kapitel benutzt – wir beschreiben die Berechnungsarten mittels Stereo- ($\tilde{\boldsymbol{F}}$-Matrix) und Trifokalgeometrie ($\tilde{\boldsymbol{T}}$-Tensor). Die mittels der Fundamentalmatrix und des Trifokaltensors gefundenen Invarianten müssen allerdings korrigiert werden, da $\tilde{\boldsymbol{F}}$ und $\tilde{\boldsymbol{T}}$ nicht mit echten Bilddaten sondern mit 3-D Punkten in der Bildebene berechnet wurden. Es wird gezeigt, wie diese Korrekturkoeffizienten gefunden werden können. Die Methoden für die Berechnung von Invarianten mittels der Fundamentalmatrix und des Trifokaltensors stammen aus der Arbeit von E.Bayro-Corrochano und J.Lasenby [14].

3.1 Geometrische projektive Invarianten

Jeder Punkt im Raum ($\mathbb{P}^3$) kann als $\boldsymbol{X_i} = \mathrm{X_i}\gamma_1 + \mathrm{Y_i}\gamma_2 + \mathrm{Z_i}\gamma_3 + \mathrm{W_i}\gamma_4$ und seine Abbildung auf die Bildebene ($\mathbb{P}^2$) als $\boldsymbol{x_i} = \mathrm{x_i}\boldsymbol{\sigma_1} + \mathrm{y_i}\boldsymbol{\sigma_2} + \mathrm{w_i}\boldsymbol{\sigma_3}$ geschrieben werden. Wählt man fünf Punkte im Raum bzw. vier Punkte in der Bildebene als projektive Basis des entsprechenden Raumes

$$\boldsymbol{X}_1 = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{X}_2 = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{X}_3 = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{X}_4 = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \end{pmatrix}, \quad \boldsymbol{X}_5 = \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix}$$

und

$$\boldsymbol{x}_1 = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{x}_3 = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \end{pmatrix}, \quad \boldsymbol{x}_4 = \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix},$$

so läßt sich jeder weitere Punkt in Termen dieser Basen folgendermaßen aufschreiben:[1]

$$\frac{\mathbf{X_i}}{\mathbf{W_i}} = \frac{[\mathbf{234}\boldsymbol{i}][\mathbf{1235}]}{[\mathbf{2345}][\mathbf{123}\boldsymbol{i}]}, \quad \frac{\mathbf{Y_i}}{\mathbf{W_i}} = \frac{[\mathbf{134}\boldsymbol{i}][\mathbf{1235}]}{[\mathbf{1345}][\mathbf{123}\boldsymbol{i}]}, \quad \frac{\mathbf{Z_i}}{\mathbf{W_i}} = \frac{[\mathbf{124}\boldsymbol{i}][\mathbf{1235}]}{[\mathbf{1245}][\mathbf{123}\boldsymbol{i}]} \tag{3.1}$$

bzw.

$$\frac{\mathbf{x_i}}{\mathbf{w_i}} = \frac{[\mathbf{23}\boldsymbol{i}][\mathbf{124}]}{[\mathbf{234}][\mathbf{12}\boldsymbol{i}]}, \quad \frac{\mathbf{y_i}}{\mathbf{w_i}} = \frac{[\mathbf{13}\boldsymbol{i}][\mathbf{124}]}{[\mathbf{134}][\mathbf{12}\boldsymbol{i}]}. \tag{3.2}$$

Bezeichnet man mit $\mathbf{(X_0, Y_0, Z_0, W_0)^T}$ Koordinaten des optischen Zentrums der Kamera, so entsteht für die Punkte im Raum und deren Abbildungen auf die Bildebene folgende geometrische Zwangsbedingung ([19]):

$$\begin{pmatrix} \mathbf{0} & \mathbf{w_5Y_5} & \mathbf{-y_5Z_5} & \mathbf{(y_5 - w_5)W_5} \\ \mathbf{w_5X_5} & \mathbf{0} & \mathbf{-x_5Z_5} & \mathbf{(x_5 - w_5)W_5} \\ \mathbf{0} & \mathbf{w_6Y_6} & \mathbf{-y_6Z_6} & \mathbf{(x_5 - w_5)W_5} \\ \mathbf{0} & \mathbf{w_6Y_6} & \mathbf{-y_6Z_6} & \mathbf{(y_6 - w_6)W_6} \\ \mathbf{w_6X_6} & \mathbf{0} & \mathbf{-x_6Z_6} & \mathbf{(x_6 - w_6)W_6} \\ \mathbf{0} & \mathbf{w_7Y_8} & \mathbf{-y_7Z_7} & \mathbf{(y_7 - w_7)W_7} \\ \mathbf{w_7X_7} & \mathbf{0} & \mathbf{-x_7Z_7} & \mathbf{(x_7 - w_7)W_7} \end{pmatrix} \begin{pmatrix} \mathbf{X_0^{-1}} \\ \mathbf{Y_0^{-1}} \\ \mathbf{Z_0^{-1}} \\ \mathbf{W_0^{-1}} \end{pmatrix} = \mathbf{0}. \tag{3.3}$$

Diese Matrix hat den Rang kleiner als 4, also ist die Determinante jeder Untermatrix, die durch vier Zeilen gebildet wird, gleich Null. Nimmt man die Determinante von der Untermatrix, die durch die ersten vier Zeilen gebildet wird, so läßt sich folgende Gleichung aufschreiben:

$$\begin{aligned} &\mathbf{(w_5y_6 - x_5y_6)X_6Z_6 + (x_5y_6 - x_5w_6)X_6W_6 +} \\ &\mathbf{(x_5w_6 - y_5w_6)X_6Y_6 + (y_5x_6 - w_5x_6)Y_6Z_6 +} \\ &\mathbf{(y_5w_6 - y_5x_6)Y_6W_6 + (w_5x_6 - w_5y_6)Z_6W_6 = 0.} \end{aligned} \tag{3.4}$$

[1]Im weiteren wird oft die Schreibweise $[\boldsymbol{abcd}]$ und $[\boldsymbol{abc}]$ statt von $[\boldsymbol{X_aX_bX_cX_d}]$ bzw. $[\boldsymbol{x_ax_bx_c}]$ benutzt, um bessere Lesbarkeit zu gewährleisten.

Obige Gleichung (3.4) kann man mit Substitution von (3.1) und (3.2) folgendermaßen umformen:

$$\begin{aligned}
&\mathbf{[125][346][1236][1246][1345][2345]} - \\
&\mathbf{[126][345][1235][1245][1346][2346]} + \\
&\mathbf{[135][246][1236][1245][1346][2345]} - \\
&\mathbf{[136][245][1235][1246][1345][2346]} + \\
&\mathbf{[145][236][1235][1246][1346][2345]} - \\
&\mathbf{[146][235][1236][1245][1345][2346]} = 0.
\end{aligned} \tag{3.5}$$

Bezeichnet man mit $\boldsymbol{i_j}$ die Koeffizienten in der Bildebene und mit $\boldsymbol{I_j}$ die Koeffizienten im Raum, so nimmt diese Gleichung folgende Gestalt an:

$$\boldsymbol{i_1 I_1 + i_2 I_2 + i_3 I_3 + i_4 I_4 + i_5 I_5 + i_6 I_6 = 0.} \tag{3.6}$$

Diese Gleichung beschreibt die Zwangsbedingungen für die Konstellation von sechs Punkten im Raum. Bestehen zwischen einigen dieser Punkte besondere geometrische Beziehungen (z.B. Kollinearität oder Koplanarität), so kann man aus (3.5) spezielle Invarianten für diese Punkte ableiten. Im Folgenden werden ein paar Beispiele für solche Invarianten gegeben.

3.1.1 4-Punkte-Koplanarität

Betrachte sechs Punkte $\boldsymbol{X_1}$,$\boldsymbol{X_2}$,$\boldsymbol{X_3}$,$\boldsymbol{X_4}$,$\boldsymbol{X_5}$ und $\boldsymbol{X_6}$. Sind vier von diesen Punkten koplanar (siehe Abbildung 3.1), so ist das entsprechende Bracket gleich Null, z.B.:

$$\boldsymbol{Inv_{cop}} = \mathbf{[1235]} = 0. \tag{3.7}$$

Den Ausdruck (3.7) kann man als die einfachste Form einer Invariante betrachten – unter beliebiger projektiver Transformation bleiben vier koplanare Punkte koplanar. Mit (3.7) vereinfacht sich (3.5) zu:

$$\begin{aligned}
&\mathbf{[125][346][1246][1345][2345]} - \\
&\mathbf{[135][246][1245][1346][2345]} + \\
&\mathbf{[146][235][1245][1345][2346]} = 0.
\end{aligned}$$

Teilen durch den mittleren Term ergibt:

$$\frac{\mathbf{[125][346][1246][1345][2345]}}{\mathbf{[135][246][1245][1346][2345]}} + \frac{\mathbf{[146][235][1245][1345][2346]}}{\mathbf{[135][246][1245][1346][2345]}} - 1 = 0.$$

Bezeichnet man mit $\mathbf{1', 2', ..., 6'}$ die Punkte in einer anderen Bildebene, so erhält man analog zur obigen Gleichung:

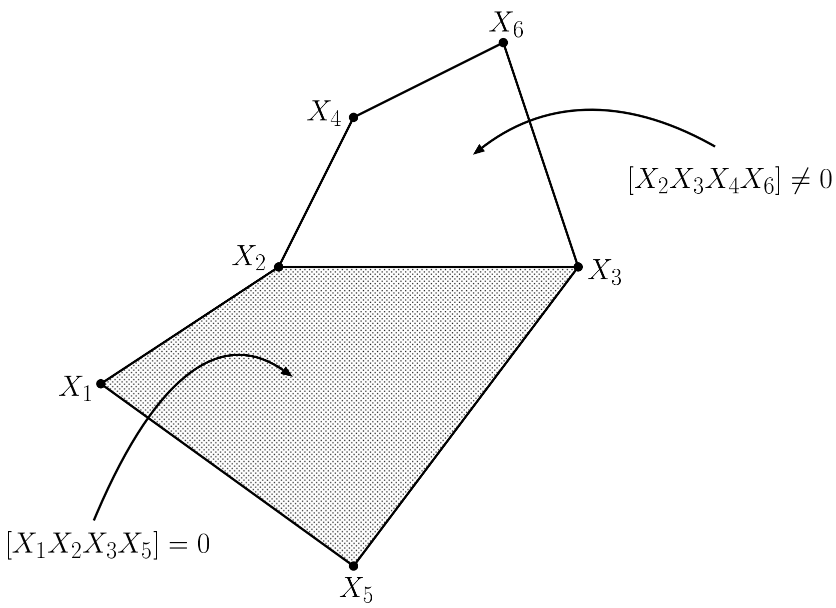


Abbildung 3.1: *6 Punkte – 1,2,3,5 sind koplanar*

$$\frac{[\mathbf{1'2'5'}][\mathbf{3'4'6'}][\mathbf{1246}][\mathbf{1345}][\mathbf{2345}]}{[\mathbf{1'3'5'}][\mathbf{2'4'6'}][\mathbf{1245}][\mathbf{1346}][\mathbf{2345}]} + \frac{[\mathbf{1'4'6'}][\mathbf{2'3'5'}][\mathbf{1245}][\mathbf{1345}][\mathbf{2346}]}{[\mathbf{1'3'5'}][\mathbf{2'4'6'}][\mathbf{1245}][\mathbf{1346}][\mathbf{2345}]} - 1 = 0.$$

Die Kombination dieser Gleichungen ergibt zwei Invarianten:

$$\boldsymbol{Inv_{cop1}} = \frac{[\mathbf{1246}][\mathbf{1345}]}{[\mathbf{1245}][\mathbf{1346}]} = \frac{[\mathbf{1'3'5'}][\mathbf{2'4'6'}][\mathbf{146}][\mathbf{235}] - [\mathbf{1'4'6'}][\mathbf{2'3'5'}][\mathbf{135}][\mathbf{246}]}{[\mathbf{1'2'5'}][\mathbf{3'4'6'}][\mathbf{146}][\mathbf{235}] - [\mathbf{1'4'6'}][\mathbf{2'3'5'}][\mathbf{125}][\mathbf{346}]} \tag{3.8}$$

und

$$\boldsymbol{Inv_{cop2}} = \frac{[\mathbf{1345}][\mathbf{2346}]}{[\mathbf{1346}][\mathbf{2345}]} = \frac{[\mathbf{1'2'5'}][\mathbf{3'4'6'}][\mathbf{135}][\mathbf{246}] - [\mathbf{1'3'5'}][\mathbf{2'4'6'}][\mathbf{125}][\mathbf{346}]}{[\mathbf{1'2'5'}][\mathbf{3'4'6'}][\mathbf{146}][\mathbf{235}] - [\mathbf{1'4'6'}][\mathbf{2'3'5'}][\mathbf{125}][\mathbf{346}]}. \tag{3.9}$$

Diese Invarianten kann man anhand der Punktkorrespondenzen in zwei Kameras berechnen.

3.1.2 Zwei 4-Punkte-Koplanaritäten

Betrachte sechs Punkte $\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}, \boldsymbol{X_4}, \boldsymbol{X_5}$ und $\boldsymbol{X_6}$. Sind die Punkte 1,2,3,5 und 2,3,4,6 koplanar (siehe Abbildung 3.2), so sind die entsprechenden Brackets gleich Null:

$$[\mathbf{1235}] = 0, \qquad [\mathbf{2346}] = 0. \tag{3.10}$$

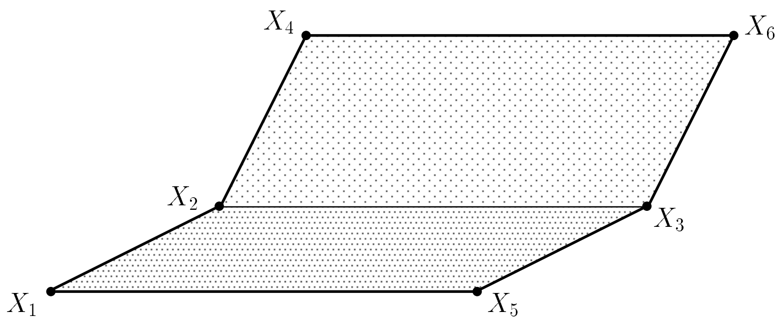


Abbildung 3.2: *6 Punkte – 1,2,3,5 und 2,3,4,6 sind koplanar*

Die Gleichung (3.5) wird zu:

$$\begin{aligned}&[\mathbf{125}][\mathbf{346}][\mathbf{1246}][\mathbf{1345}] - \\ &[\mathbf{135}][\mathbf{246}][\mathbf{1245}][\mathbf{1346}] = 0.\end{aligned}$$

Daraus ergibt sich folgende Invariante:

$$\boldsymbol{Inv_{2cop}} = \frac{[\mathbf{1245}][\mathbf{1346}]}{[\mathbf{1246}][\mathbf{1345}]} = \frac{[\mathbf{125}][\mathbf{346}]}{[\mathbf{135}][\mathbf{246}]}. \tag{3.11}$$

Diese Gleichung demonstriert, daß mit der Zunahme der einschränkenden Bedingungen an die Punkte die entsprechenden Invarianten weniger Informationen zur Berechnung benötigen – in diesem Fall reicht schon eine Kamera, um die Invariante abzuschätzen.

3.2 Berechnung von Invarianten mit zwei Kameras ($\boldsymbol{F}$)

Im Abschnitt 2.2 wurde die geometrische Interpretation der $\tilde{\boldsymbol{F}}$-Matrix gegeben. Daraus läßt sich ableiten, wie man die Rauminvarianten in der Bildebene mittels $\tilde{\boldsymbol{F}}$-Matrix berechnet.

Betrachte zwei Punke im Raum $\boldsymbol{X_1}$ und $\boldsymbol{X_2}$. Seien $\boldsymbol{A'_1}$ und $\boldsymbol{B'_1}$ bzw. $\boldsymbol{A'_2}$ und $\boldsymbol{B'_2}$ die Abbildungen dieser Punkte in die Kameras $\boldsymbol{A}$ und $\boldsymbol{B}$ mit den optischen Zentren $\boldsymbol{A_0}$ und $\boldsymbol{B_0}$. Weiterhin sei $\boldsymbol{L_{12}}$ die Gerade, die die Punkte $\boldsymbol{X_1}$ und $\boldsymbol{X_2}$ verbindet, und $\boldsymbol{L^A_{12}}$,$\boldsymbol{L^B_{12}}$ ihre Abbildungen in den Kameras. Die optischen Ebenen, die $\boldsymbol{L^A_{12}}$ und $\boldsymbol{L^B_{12}}$ enthalten, schneiden sich entlang der Gerade $\boldsymbol{L_{12}}$:

$$\boldsymbol{L_{12}} = \boldsymbol{X_1} \bigwedge \boldsymbol{X_2} = (\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{12}}) \bigvee (\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{12}}).$$

Analog stellt man die Gleichung für die Punkte $\boldsymbol{X_3}$ und $\boldsymbol{X_4}$ auf:

$$\boldsymbol{L_{34}} = \boldsymbol{X_3} \bigwedge \boldsymbol{X_4} = (\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{34}}) \bigvee (\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{34}}).$$

Mit obigen Gleichungen kann man das Bracket $[\mathbf{1234}]$ folgendermaßen berechnen:[2]

$$\begin{aligned}
[\mathbf{1234}] &= (\boldsymbol{X_1} \bigwedge \boldsymbol{X_2} \bigwedge \boldsymbol{X_3} \bigwedge \boldsymbol{X_4})\boldsymbol{\mathcal{I}^{-1}} \\
&= \{((\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{12}}) \bigvee (\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{12}})) \bigwedge ((\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{34}}) \bigvee (\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{34}}))\}\boldsymbol{\mathcal{I}^{-1}} \\
&= -\{((\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{12}}) \bigvee (\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{34}})) \bigwedge ((\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{12}}) \bigvee (\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{34}}))\}\boldsymbol{\mathcal{I}^{-1}} \\
&= -\{(\boldsymbol{A_0} \bigwedge \boldsymbol{A'_{1234}}) \bigwedge (\boldsymbol{B_0} \bigwedge \boldsymbol{B'_{1234}})\}\boldsymbol{\mathcal{I}^{-1}} \\
&= (\boldsymbol{A_0} \bigwedge \boldsymbol{B_0} \bigwedge \boldsymbol{A'_{1234}} \bigwedge \boldsymbol{B'_{1234}})\boldsymbol{\mathcal{I}^{-1}}, \qquad (3.12)
\end{aligned}$$

wobei $\boldsymbol{A'_{1234}}$ und $\boldsymbol{B'_{1234}}$ die Schnittpunke der Geraden $\boldsymbol{L^A_{12}}$ und $\boldsymbol{L^A_{34}}$ bzw. $\boldsymbol{L^B_{12}}$ und $\boldsymbol{L^B_{34}}$ sind (siehe Abbildung 3.3).

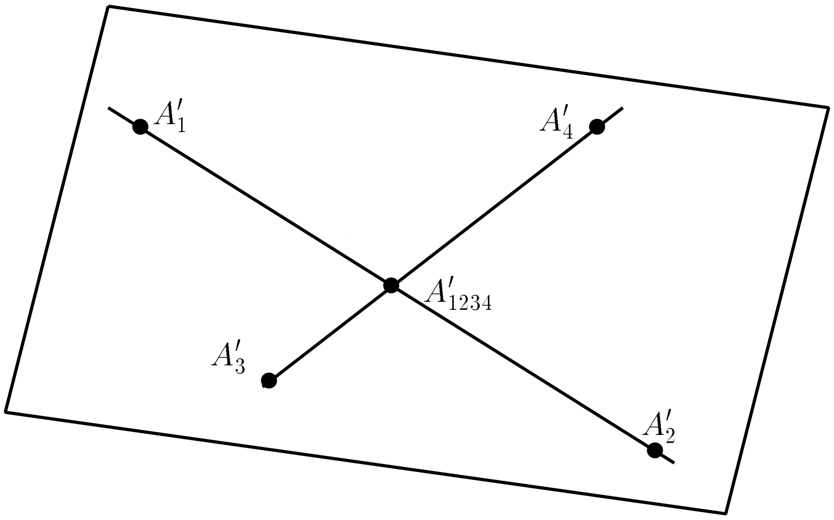


Abbildung 3.3: *Punkt* $\boldsymbol{A'_{1234}}$ *ist der Schnittpunkt der Geraden* $\boldsymbol{A'_1} \wedge \boldsymbol{A'_2}$ *und* $\boldsymbol{A'_3} \wedge \boldsymbol{A'_4}$ *in der Bildebene*

Drückt man die Schnittpunkte in Termen von Basisvektoren der Kameras aus

$$\begin{aligned}
\boldsymbol{A'_{1234}} &= \boldsymbol{\alpha_{1234_1}A_1} + \boldsymbol{\alpha_{1234_2}A_2} + \boldsymbol{\alpha_{1234_3}A_3}, \\
\boldsymbol{B'_{1234}} &= \boldsymbol{\beta_{1234_1}B_1} + \boldsymbol{\beta_{1234_2}B_2} + \boldsymbol{\beta_{1234_3}B_3},
\end{aligned}$$

so läßt sich obige Gleichung mit Hilfe von (2.24) folgendermaßen aufschreiben:

$$\begin{aligned}
[\mathbf{1234}] = (\boldsymbol{X_1} \bigwedge \boldsymbol{X_2} \bigwedge \boldsymbol{X_3} \bigwedge \boldsymbol{X_4})\boldsymbol{\mathcal{I}^{-1}} &= (\boldsymbol{A_0} \bigwedge \boldsymbol{B_0} \bigwedge \boldsymbol{A'_{1234}} \bigwedge \boldsymbol{B'_{1234}})\boldsymbol{\mathcal{I}^{-1}} \\
&= \sum_{\boldsymbol{i,j=1}}^{\mathbf{3}} \boldsymbol{\tilde{F}_{ij}\alpha_{1234_i}\beta_{1234_j}}. \qquad (3.13)
\end{aligned}$$

[2]Folgt aus der Gleichung $(\boldsymbol{\Phi_1} \bigvee \boldsymbol{\Phi_2}) \bigwedge (\boldsymbol{\Phi_3} \bigvee \boldsymbol{\Phi_4}) = -(\boldsymbol{\Phi_1} \bigvee \boldsymbol{\Phi_3}) \bigwedge (\boldsymbol{\Phi_2} \bigvee \boldsymbol{\Phi_4})$. Für den Beweis siehe Anhang B.

Ist erst ein Bracket in Termen der Fundamentalmatrix gefunden, dann kann man auch die Invarianten in diesen Termen bestimmen, z.B.:

$$\boldsymbol{Inv_{2cop}} = \frac{[\mathbf{1245}][\mathbf{1346}]}{[\mathbf{1246}][\mathbf{1345}]} = \frac{(\boldsymbol{\beta_{1245}\tilde{F}\alpha^T_{1245}})(\boldsymbol{\beta_{1346}\tilde{F}\alpha^T_{1346}})}{(\boldsymbol{\beta_{1246}\tilde{F}\alpha^T_{1246}})(\boldsymbol{\beta_{1345}\tilde{F}\alpha^T_{1345}})}, \tag{3.14}$$

wobei $\boldsymbol{\alpha_{ijkl}} = (\boldsymbol{\alpha_{ijkl_1}}, \boldsymbol{\alpha_{ijkl_2}}, \boldsymbol{\alpha_{ijkl_3}})$ und $\boldsymbol{\beta_{ijkl}} = (\boldsymbol{\beta_{ijkl_1}}, \boldsymbol{\beta_{ijkl_2}}, \boldsymbol{\beta_{ijkl_3}})$ sind.

Allerdings sind die Invarianten, die nach (3.14) mit echten Daten berechnet werden, zu korrigieren, da in obigen Gleichungen keine echten Bildpunkte sondern die Punkte im Raum benutzt wurden. Um die Invarianten auch mit Bildpunkten korrekt zu berechnen, braucht man Korrekturkoeffizienten. Diese werden auf folgende Weise berechnet ([14]):

> Betrachte einen Punkt $\boldsymbol{X}$ im Raum. Sei $\boldsymbol{a'}$ die $\mathbb{P}^2$-Projektion dieses Punktes in die Bildebene der Kamera $\boldsymbol{A}$ und $\boldsymbol{A'}$ die $\mathbb{P}^3$-Repräsentation von $\boldsymbol{a'}$. Der Bildpunkt $\boldsymbol{a'}$ kann in Termen der $\mathbb{P}^2$-Bildebenenbasis folgendermaßen geschrieben werden:
>
> $$\boldsymbol{a'} = \boldsymbol{a_1} + \boldsymbol{\delta_1}(\boldsymbol{a_2} - \boldsymbol{a_1}) + \boldsymbol{\delta_2}(\boldsymbol{a_3} - \boldsymbol{a_1}) = \boldsymbol{\lambda^A_1 a_1} + \boldsymbol{\lambda^A_2 a_2} + \boldsymbol{\lambda^A_3 a_3}.$$
>
> Andererseits kann man $\boldsymbol{a'}$ mit Hilfe des projektiven splits von $\boldsymbol{A'}$ ableiten:
>
> $$\boldsymbol{a'} = \frac{\boldsymbol{A'} \wedge \boldsymbol{\gamma_4}}{\boldsymbol{A'} \cdot \boldsymbol{\gamma_4}} = \frac{\boldsymbol{\alpha_1}(\boldsymbol{A_1} \wedge \boldsymbol{\gamma_4}) + \boldsymbol{\alpha_2}(\boldsymbol{A_2} \wedge \boldsymbol{\gamma_4}) + \boldsymbol{\alpha_3}(\boldsymbol{A_3} \wedge \boldsymbol{\gamma_4})}{\boldsymbol{A'} \cdot \boldsymbol{\gamma_4}}.$$
>
> Somit ergibt sich folgende Relation zwischen den $\boldsymbol{\alpha_i}$'s und $\boldsymbol{\lambda^A_i}$'s:
>
> $$\boldsymbol{\alpha_i} = \frac{\boldsymbol{A'} \cdot \boldsymbol{\gamma_4}}{\boldsymbol{A_i} \cdot \boldsymbol{\gamma_4}} \boldsymbol{\lambda^A_i}. \tag{3.15}$$
>
> Analog zur Kamera $\boldsymbol{A}$ sei $\boldsymbol{b'} = \boldsymbol{\lambda^B_1 b_1} + \boldsymbol{\lambda^B_2 b_2} + \boldsymbol{\lambda^B_3 b_3}$ die $\mathbb{P}^2$-Projektion von $\boldsymbol{X}$ in Kamera $\boldsymbol{B}$. Ist $\boldsymbol{B'} = \sum_{i=1}^{3} \boldsymbol{\beta_i B_i}$ die $\mathbb{P}^3$-Repräsentation von $\boldsymbol{b'}$, so gilt:
>
> $$\boldsymbol{\beta_i} = \frac{\boldsymbol{B'} \cdot \boldsymbol{\gamma_4}}{\boldsymbol{B_i} \cdot \boldsymbol{\gamma_4}} \boldsymbol{\lambda^B_i}. \tag{3.16}$$

Mit echten Daten (Punktkorrespondenzen $\boldsymbol{a'_i} = (\boldsymbol{\lambda^A_{i1}}, \boldsymbol{\lambda^A_{i2}}, \boldsymbol{\lambda^A_{i3}})^{\boldsymbol{T}}$ und $\boldsymbol{b'_i} = (\boldsymbol{\lambda^B_{i1}}, \boldsymbol{\lambda^B_{i2}}, \boldsymbol{\lambda^B_{i3}})^{\boldsymbol{T}}$ zwischen den Kameras $\boldsymbol{A}$ und $\boldsymbol{B}$) läßt sich die 'echte' $\boldsymbol{F}$-Matrix bestimmen. Diese Matrix genügt der Gleichung

$$\boldsymbol{b'^T_i F a'_i} = \mathbf{0} \tag{3.17}$$

und kann z.B. mit dem 8-Punkte-Algorithmus berechnet werden.

Definiert man

$$\tilde{\boldsymbol{F}}_{kl} = (\boldsymbol{A}_k \cdot \gamma_4)(\boldsymbol{B}_l \cdot \gamma_4)\boldsymbol{F}_{kl},$$

so ergibt sich mit (3.15) und (3.16) folgende Beziehung:

$$\tilde{\boldsymbol{F}}_{kl}\boldsymbol{\alpha}_{ik}\boldsymbol{\beta}_{il} = (\boldsymbol{A}'_i \cdot \gamma_4)(\boldsymbol{B}'_i \cdot \gamma_4)\boldsymbol{F}_{kl}\boldsymbol{\lambda}^{A}_{ik}\boldsymbol{\lambda}^{B}_{il}. \tag{3.18}$$

Somit kann man die Gleichung (3.14) in Termen der 'echten' $\boldsymbol{F}$-Matrix umschreiben:

$$\boldsymbol{Inv_{2cop}} = \frac{(\boldsymbol{b}'^T_{1245}\boldsymbol{F}\boldsymbol{a}'_{1245})(\boldsymbol{b}'^T_{1346}\boldsymbol{F}\boldsymbol{a}'_{1346})}{(\boldsymbol{b}'^T_{1246}\boldsymbol{F}\boldsymbol{a}'_{1246})(\boldsymbol{b}'^T_{1345}\boldsymbol{F}\boldsymbol{a}'_{1345})}\frac{\boldsymbol{\phi}_{1245}\boldsymbol{\phi}_{1346}}{\boldsymbol{\phi}_{1246}\boldsymbol{\phi}_{1345}}, \tag{3.19}$$

wobei

$$\boldsymbol{\phi}_{ijkl} = (\boldsymbol{A}'_{ijkl} \cdot \gamma_4)(\boldsymbol{B}'_{ijkl} \cdot \gamma_4). \tag{3.20}$$

Werden $\boldsymbol{\mu}_{ijkl}$ und $\boldsymbol{\nu}_{ijkl}$ durch

$$\begin{aligned} \boldsymbol{a}'_{ijkl} &= \boldsymbol{\mu}_{ijkl}\boldsymbol{a}'_l + (\boldsymbol{1} - \boldsymbol{\mu}_{ijkl})\boldsymbol{a}'_k, \\ \boldsymbol{b}'_{ijkl} &= \boldsymbol{\nu}_{ijkl}\boldsymbol{b}'_l + (\boldsymbol{1} - \boldsymbol{\nu}_{ijkl})\boldsymbol{b}'_k \end{aligned}$$

definiert, so gilt für $\boldsymbol{Inv_{2cop}}$:[3]

$$\frac{\boldsymbol{\phi}_{1245}\boldsymbol{\phi}_{1346}}{\boldsymbol{\phi}_{1246}\boldsymbol{\phi}_{1345}} = \frac{\boldsymbol{\mu}_{1245}\boldsymbol{\mu}_{1346}\boldsymbol{\nu}_{1245}\boldsymbol{\nu}_{1346}}{\boldsymbol{\mu}_{1246}\boldsymbol{\mu}_{1345}\boldsymbol{\nu}_{1246}\boldsymbol{\nu}_{1345}}.$$

Also ist

$$\boldsymbol{Inv_{2cop}} = \frac{(\boldsymbol{b}'^T_{1245}\boldsymbol{F}\boldsymbol{a}'_{1245})(\boldsymbol{b}'^T_{1346}\boldsymbol{F}\boldsymbol{a}'_{1346})}{(\boldsymbol{b}'^T_{1246}\boldsymbol{F}\boldsymbol{a}'_{1246})(\boldsymbol{b}'^T_{1345}\boldsymbol{F}\boldsymbol{a}'_{1345})}\frac{\boldsymbol{\mu}_{1245}\boldsymbol{\mu}_{1346}\boldsymbol{\nu}_{1245}\boldsymbol{\nu}_{1346}}{\boldsymbol{\mu}_{1246}\boldsymbol{\mu}_{1345}\boldsymbol{\nu}_{1246}\boldsymbol{\nu}_{1345}} \tag{3.21}$$

nur in Termen der betrachteten Bildpunkte ausgedrückt, d.h. es ist möglich, rein aus den Punktkorrespondenzen von zwei Kameras Invarianten zu berechnen.

[3]Siehe Anhang C.

3.3 Berechnung von Invarianten mit drei Kameras ($\boldsymbol{T}$)

Das Konzept der Berechnung von Invarianten mit zwei Kameras kann man auf drei Kameras erweitern:

Betrachte vier Punkte $\boldsymbol{X_1},\boldsymbol{X_2},\boldsymbol{X_3},\boldsymbol{X_4}$ und zwei Geraden $\boldsymbol{L_{12}} = \boldsymbol{X_1} \bigwedge \boldsymbol{X_2}$ und $\boldsymbol{L_{34}} = \boldsymbol{X_3} \bigwedge \boldsymbol{X_4}$ im Raum. Seien $\boldsymbol{A'_1},\boldsymbol{A'_2},\boldsymbol{A'_3},\boldsymbol{A'_4},\boldsymbol{L^A_{12}},\boldsymbol{L^A_{34}}$, $\boldsymbol{B'_1},\boldsymbol{B'_2},\boldsymbol{B'_3},\boldsymbol{B'_4},\boldsymbol{L^B_{12}},\boldsymbol{L^B_{34}}$, $\boldsymbol{C'_1},\boldsymbol{C'_2},\boldsymbol{C'_3},\boldsymbol{C'_4},\boldsymbol{L^C_{12}},\boldsymbol{L^C_{34}}$ deren Abbildungen in den Kameras $\boldsymbol{A}$,$\boldsymbol{B}$ und $\boldsymbol{C}$. Analog zum Abschnitt 3.2 (und mit gleichen Basen für Vektoren und Bivektoren in der Bildebene) kann man folgende Gleichungen aufstellen:

$$\begin{aligned}
\boldsymbol{X_1} \bigwedge \boldsymbol{X_2} &= (\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{12}}) \bigvee (\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{12}}),\\
\boldsymbol{X_3} \bigwedge \boldsymbol{X_4} &= (\boldsymbol{A_0} \bigwedge \boldsymbol{L^A_{34}}) \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L^C_{34}}).
\end{aligned}$$

Daraus ergibt sich folgende Gleichung für das Bracket $[\mathbf{1234}]$:

$$\begin{aligned}
\{\boldsymbol{X_1} \bigwedge \boldsymbol{X_2} \bigwedge \boldsymbol{X_3} \bigwedge \boldsymbol{X_4}\}\boldsymbol{\mathcal{I}^{-1}} &= \{(\boldsymbol{A_0} \bigwedge \boldsymbol{A'_{1234}}) \bigwedge ((\boldsymbol{B_0} \bigwedge \boldsymbol{L^B_{12}})\\
&\qquad \bigvee (\boldsymbol{C_0} \bigwedge \boldsymbol{L^C_{34}}))\}\boldsymbol{\mathcal{I}^{-1}}.
\end{aligned} \tag{3.22}$$

Vergleicht man (2.29) mit der obigen Gleichung, so folgt:

$$\{\boldsymbol{X_1} \bigwedge \boldsymbol{X_2} \bigwedge \boldsymbol{X_3} \bigwedge \boldsymbol{X_4}\}\boldsymbol{\mathcal{I}^{-1}} = \boldsymbol{\tilde{T}\alpha_{1234} l^B_{12} l^C_{34}}, \tag{3.23}$$

wobei $\boldsymbol{A'_{1234}} = \sum_{i=1}^{3} \boldsymbol{\alpha_{1234_i} A_i}$, $\boldsymbol{L^B_{12}} = \sum_{i=1}^{3} \boldsymbol{l^B_{12} L^B_i}$ und $\boldsymbol{L^C_{34}} = \sum_{i=1}^{3} \boldsymbol{l^C_{34} L^C_i}$ sind.

Kann man ein Bracket in Termen des Trifokaltensors berechnen, so lassen sich auch Invarianten in diesen Termen ausdrücken, z.B.:

$$\boldsymbol{Inv_{2cop}} = \frac{[\mathbf{1245}][\mathbf{1346}]}{[\mathbf{1246}][\mathbf{1345}]} = \frac{(\boldsymbol{\tilde{T}\alpha_{1245} l^B_{12} l^C_{45}})(\boldsymbol{\tilde{T}\alpha_{1346} l^B_{13} l^C_{46}})}{(\boldsymbol{\tilde{T}\alpha_{1246} l^B_{12} l^C_{46}})(\boldsymbol{\tilde{T}\alpha_{1345} l^B_{13} l^C_{45}})}. \tag{3.24}$$

Will man Invarianten mit echten Daten berechnen, so ist es notwendig, die Korrekturkoeffizienten wie im Fall der Fundamentalmatrix zu berechnen:

Korrekturkoeffizienten für Punkte sind im Abschnitt 3.2 berechnet:

$$\boldsymbol{\alpha_i} = \frac{\boldsymbol{A'} \cdot \gamma_4}{\boldsymbol{A_i} \cdot \gamma_4} \boldsymbol{\lambda_i^A}. \tag{3.25}$$

Korrekturkoeffizienten für Geraden kann man von Punkten ableiten – so gilt z.B. für Kamera $\boldsymbol{B}$:

$$\boldsymbol{l^B_{pq_k}} = \{\boldsymbol{\beta_p} \times \boldsymbol{\beta_q}\}_k = \boldsymbol{\epsilon_{kij}\beta_{p_i}\beta_{q_j}} = \frac{(\boldsymbol{B'_p} \cdot \gamma_4)(\boldsymbol{B'_q} \cdot \gamma_4)}{(\boldsymbol{B_i} \cdot \gamma_4)(\boldsymbol{B_j} \cdot \gamma_4)} \boldsymbol{\lambda^B_{p_i}\lambda^B_{q_j}}, \tag{3.26}$$

wobei $\times$ das Standardvektorprodukt ist. ϵ_{kij} ist $+1$ oder -1, wenn kij eine gerade bzw. ungerade Permutation von 123 ist, und sonst 0.

Analog zum Abschnitt 3.2 erhält man:

$$\boldsymbol{Inv_{2cop}} = \frac{[\mathbf{1245}][\mathbf{1346}]}{[\mathbf{1246}][\mathbf{1345}]} = \frac{(\boldsymbol{Ta'_{1245}l^B_{12}l^C_{45}})(\boldsymbol{Ta'_{1346}l^B_{13}l^C_{46}})}{(\boldsymbol{Ta'_{1246}l^B_{12}l^C_{46}})(\boldsymbol{Ta'_{1345}l^B_{13}l^C_{45}})} \frac{\boldsymbol{\psi_{1245}\psi_{1346}}}{\boldsymbol{\psi_{1246}\psi_{1345}}}, \tag{3.27}$$

mit

$$\boldsymbol{\psi_{ijkl}} = \boldsymbol{A'_{ijkl}} \cdot \gamma_4.$$

Sind die $\boldsymbol{\mu}$'s wie im Abschnitt 3.2 definiert, so ergibt sich folgende Gleichung für $\boldsymbol{Inv_{2cop}}$:

$$\boldsymbol{Inv_{2cop}} = \frac{[\mathbf{1245}][\mathbf{1346}]}{[\mathbf{1246}][\mathbf{1345}]} = \frac{(\boldsymbol{Ta'_{1245}l^B_{12}l^C_{45}})(\boldsymbol{Ta'_{1346}l^B_{13}l^C_{46}})}{(\boldsymbol{Ta'_{1246}l^B_{12}l^C_{46}})(\boldsymbol{Ta'_{1345}l^B_{13}l^C_{45}})} \frac{\boldsymbol{\mu_{1245}\mu_{1346}}}{\boldsymbol{\mu_{1246}\mu_{1345}}}. \tag{3.28}$$

Die Invariante $\boldsymbol{Inv_{2cop}}$ wurde also nur mit den Punktkorrespondenzen in drei Kameras berechnet.

Kapitel 4

Rekonstruktion der Gestalt und Bewegung

Mit den theoretischen Grundlagen, die in vorhergegangenen Kapiteln behandelt wurden, kann man verschiedende Probleme lösen. Hier werden diese Kenntnisse eingesetzt, um die Szenenrekonstruktion durchzuführen. Unter einer *Szene* verstehen wir eine Konstellation von Objekten im Raum, die mit mehreren Kameras aufgenommen wird. Die gemessenen Daten sind dabei die Abbildungen von diesen Objekten in den Kameras. Nur anhand dieser Meßdaten wird die 3-D Rekonstruktion von Objekten (Gestalt) und von den Kamerapositionen (Bewegung) gemacht. In dem Ansatz, den wir verfolgen, wird nur die Punktrekonstruktion durchgeführt. Dabei gehen wir davon aus, daß die Punktkorrespondenzen in den Bildern schon bestimmt sind.

Am Anfang werden verschiedene Möglichkeiten der projektiven Rekonstruktion gezeigt. Die projektive Rekonstruktion bedeutet in diesem Zusammenhang eine Rekonstruktion bis auf eine projektive Transformation, d.h. es gibt eine projektive Transformation, die die rekonstruierten Objekte in die euklidische Basis des Raumes überführt. Weiterhin wird gezeigt, wie man anhand des a priori Wissens über die Szene solche Transformationen finden kann, sowie die verdeckten Punkte von Objekten rekonstruieren und durch die Analyse von Bildern die Szenenpunkte ergänzen kann.

4.1 Projektive Rekonstruktion mittels SVD

Zunächst wollen wir eine Methode zur projektiven Rekonstruktion vorstellen, die auf dem Singulärwertzerlegung-Algorithmus (SVD) basiert. Ursprünglich wurde die SVD-Methode von C.Tomasi und T.Kanade für die euklidische Rekonstruktion entwickelt ([4]). W.Triggs und P.Sturm haben diese Methode auf projektive Rekonstruktion verallgemeinert ([15]).

Diese Art der Rekonstruktion erfordert Kentnisse über die *projektive Tiefe* – die Beziehung zwischen dem Punkt im Raum und deren Abbildung in der Kamera. Mit diesen Tiefen wird dann das *Join-Bild* erzeugt, das anschließend in Gestalt- und

Bewegungsanteile mit Hilfe der Singulärwertzerlegung gespaltet wird.

Im Folgenden werden die Begriffe projektive Tiefe und Join-Bild näher erläutert. Es werden die Möglichkeiten zur Berechnung der projektiven Tiefe mittels Epipolargeometrie und Invarianten erklärt. Das Join-Bild wird definiert. Anschließend wird der SVD-Algorithmus beschrieben.

4.1.1 Projektive Tiefe

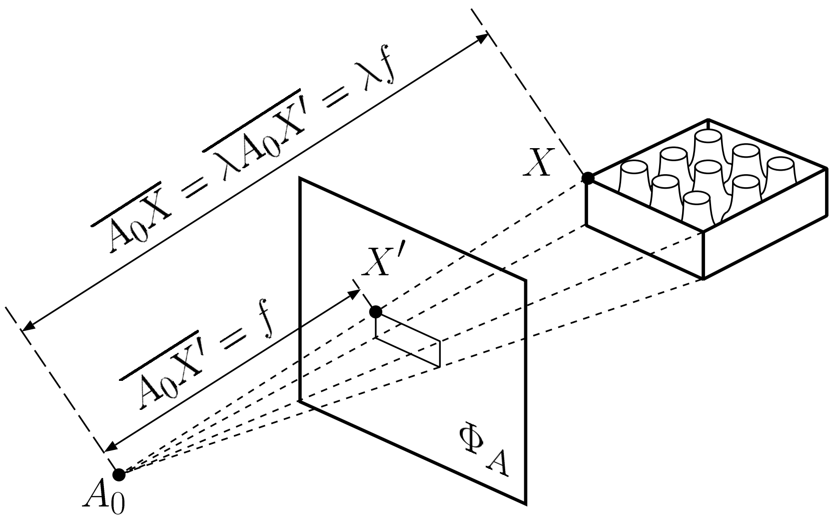


Abbildung 4.1: *Geometrische Interpretation der projektiven Tiefe*

Geometrisch gesehen beschreibt die *projektive Tiefe* die Beziehung zwischen dem Abstand des Punktes $\boldsymbol{X}$ aus $\mathbb{P}^3$ und dem optischen Zentrum der Kamera und der Fokallänge $\boldsymbol{f}$, wie es in der Abbildung 4.1 dargestellt ist.

Die projektive Tiefe läßt sich auf algebraische Weise von der Projektionsgleichung

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{X}$$

auf folgende Weise ableiten:

$$\boldsymbol{x} = \boldsymbol{\lambda} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{1} \end{bmatrix} = \boldsymbol{P}\boldsymbol{X} = \boldsymbol{P} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \\ \mathbf{1} \end{bmatrix}. \tag{4.1}$$

Der skalare Faktor $\boldsymbol{\lambda}$ ist dabei die gesuchte projektive Tiefe.

Die praktische Berechnung der projektiven Tiefe aus der Gleichung (4.1) läßt sich auf zwei verschiedene Weisen durchführen.

Berechnung der Tiefe mittels Fundamentalmatrix

Betrachte zwei Kameras $\boldsymbol{A}$ und $\boldsymbol{B}$ und einen Punkt $\boldsymbol{X}$ im Raum. Seien $\boldsymbol{x_A}$ und $\boldsymbol{x_B}$ die Abbildungen von $\boldsymbol{X}$ in $\boldsymbol{A}$ und $\boldsymbol{B}$, $\boldsymbol{\lambda_A}$,$\boldsymbol{\lambda_B}$ die projektiven Tiefen dieser Punkte bezüglich $\boldsymbol{A}$ bzw. $\boldsymbol{B}$, $\boldsymbol{e_{AB}}$ der Epipol von Kamera $\boldsymbol{A}$ in der Kamera $\boldsymbol{B}$ (Abbildung des optischen Zentrums von $\boldsymbol{A}$ in $\boldsymbol{B}$) und $\boldsymbol{F_{AB}}$ die Fundamentalmatrix von $\boldsymbol{A}$ nach $\boldsymbol{B}$. Weiterhin sei $\boldsymbol{l_B^X}$ die Epipolargerade von Punkt $\boldsymbol{X}$ in Kamera $\boldsymbol{B}$. Diese Gerade kann man auf zwei verschiedene Arten berechen.

Mittels Fundamentalmatrix:

$$\boldsymbol{l_B^X} = \boldsymbol{F_{AB}}(\boldsymbol{\lambda_A x_A}), \tag{4.2}$$

oder als join von dem Epipol $\boldsymbol{e_{AB}}$ mit dem Bildpunkt $\boldsymbol{x_B}$:

$$\boldsymbol{l_B^X} = \boldsymbol{e_{AB}} \bigwedge (\boldsymbol{\lambda_B x_B}). \tag{4.3}$$

Setzt man die rechten Seiten von (4.2) und (4.3) gleich

$$\boldsymbol{e_{AB}} \bigwedge (\boldsymbol{\lambda_B x_B}) = \boldsymbol{F_{AB}}(\boldsymbol{\lambda_A x_A}),$$

so ergibt sich folgende Beziehung zwischen den projektiven Tiefen des Punktes $\boldsymbol{X}$ bezüglich den Kameras $\boldsymbol{A}$ und $\boldsymbol{B}$ ([15]):

$$\boldsymbol{\lambda_B} = \frac{(\boldsymbol{e_{AB}} \bigwedge \boldsymbol{x_B}) \cdot (\boldsymbol{F_{AB} x_A})}{||\boldsymbol{e_{AB}} \bigwedge \boldsymbol{x_B}||^2} \boldsymbol{\lambda_A}. \tag{4.4}$$

Dies ist die Gleichung für die Berechnung von projektiven Tiefen mittels Fundamentalmatrix anhand von Meßdaten.

Berechnung der Tiefe mittels Invarianten

Sei $\boldsymbol{X}$ ein Punkt im $\mathbb{P}^3$ und $\boldsymbol{x_A}$,$\boldsymbol{x_B}$,$\boldsymbol{x_C}$,$\boldsymbol{x_D}$ die Projektionen dieses Punktes in den Kameras $\boldsymbol{A}$,$\boldsymbol{B}$,$\boldsymbol{C}$ und $\boldsymbol{D}$ mit den optischen Zentren $\boldsymbol{A_0}$,$\boldsymbol{B_0}$,$\boldsymbol{C_0}$,$\boldsymbol{D_0}$. Laut (4.1) ist die projektive Tiefe von $\boldsymbol{X}$ bezüglich $\boldsymbol{A}$

$$\boldsymbol{\lambda_A x_A} = \boldsymbol{P_A X}. \tag{4.5}$$

Ist der Weltkoordinatenursprung in dem optischen Zentrum von $\boldsymbol{A}$, so gilt:[1]

[1]Siehe Anhang A.

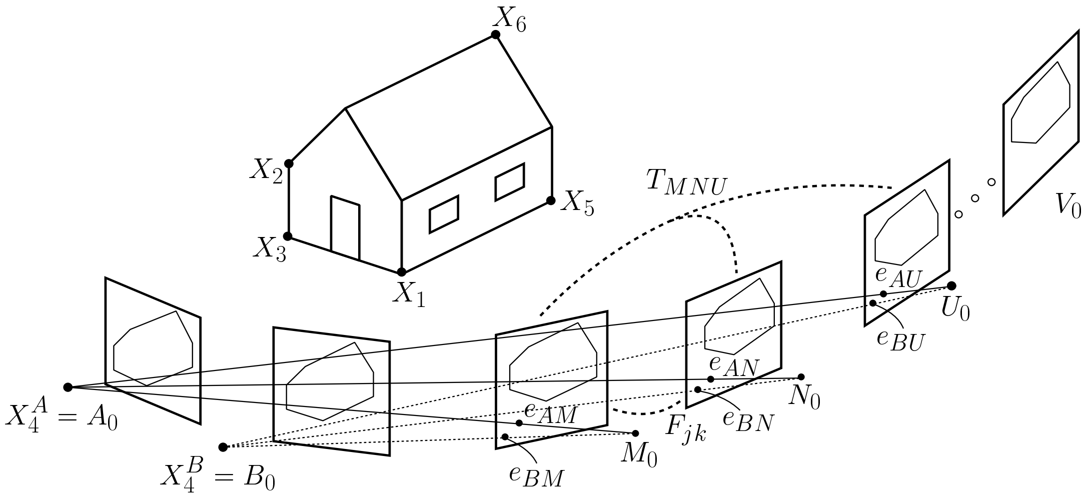


Abbildung 4.2: *Berechnung der projektiven Tiefe mittels Invarianten*

$$\boldsymbol{\lambda_A x_A} = \begin{bmatrix} \boldsymbol{f} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{f} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \\ \mathbf{1} \end{bmatrix}. \tag{4.6}$$

Daraus ergibt sich:

$$\boldsymbol{\lambda_A} = \mathbf{Z}. \tag{4.7}$$

Diese Gleichung besagt, daß die projektive Tiefe eines Punktes nur von der **Z**-Komponente dieses Punktes abhängig ist.

Wählt man im Raum vier weitere Punkte in allgemeiner Position[2], die zusammen mit $\boldsymbol{A_0}$ eine projektive Basis des $\mathbb{P}^3$-Raumes bilden

$$\boldsymbol{X_1} = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \ \boldsymbol{X_2} = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \ \boldsymbol{X_3} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \end{pmatrix}, \ \boldsymbol{X_4} = \boldsymbol{A_0} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \end{pmatrix}, \ \boldsymbol{X_5} = \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{pmatrix},$$

so kann man (4.7) mit Hilfe von (3.1) in Termen von Invarianten ausdrücken:

$$\boldsymbol{\lambda_A} = \mathbf{Z} = \frac{[\boldsymbol{X_1 X_2 A_0 X}][\boldsymbol{X_1 X_2 X_3 X_5}]}{[\boldsymbol{X_1 X_2 A_0 X_5}][\boldsymbol{X_1 X_2 X_3 X}]}. \tag{4.8}$$

Diese Invariante kann man auch nur mit gemessenen Daten in zwei (bei $\boldsymbol{F}$) oder drei (bei $\boldsymbol{T}$) Bildern berechnen, wenn berücksichtigt wird, daß die Abbildungen des

[2]siehe Anhang E.

optischen Zentrums $\boldsymbol{A_0}$ in die Kameras $\boldsymbol{B}$, $\boldsymbol{C}$,$\boldsymbol{D}$ die Epipole $\boldsymbol{e_{AB}}$,$\boldsymbol{e_{AC}}$ und $\boldsymbol{e_{AD}}$ sind. So ergibt sich mit (3.19)

$$\boldsymbol{\lambda_A} = \frac{\boldsymbol{U_1 U_2}}{\boldsymbol{U_3 U_4}} \frac{\boldsymbol{\phi_1 \phi_2}}{\boldsymbol{\phi_3 \phi_4}}, \tag{4.9}$$

wobei

$$\begin{aligned}
\boldsymbol{U_1} &= \boldsymbol{x^T_{C_{(X_1,X_2,E_{AC},X)}} F_{BC} x_{B_{(X_1,X_2,E_{AB},X)}}},\\
\boldsymbol{U_2} &= \boldsymbol{x^T_{C_{(X_1,X_2,X_3,X_5)}} F_{BC} x_{B_{(X_1,X_2,X_3,X_5)}}},\\
\boldsymbol{U_3} &= \boldsymbol{x^T_{C_{(X_1,X_2,E_{AC},X_5)}} F_{BC} x_{B_{(X_1,X_2,E_{AB},X_5)}}},\\
\boldsymbol{U_4} &= \boldsymbol{x^T_{C_{(X_1,X_2,X_3,X)}} F_{BC} x_{B_{(X_1,X_2,X_3,X)}}},\\
\boldsymbol{\phi_1} &= \boldsymbol{\mu_{B_{(X_1,X_2,E_{AB},X)}} \nu_{C_{(X_1,X_2,E_{AC},X)}}},\\
\boldsymbol{\phi_2} &= \boldsymbol{\mu_{B_{(X_1,X_2,X_3,X_5)}} \nu_{C_{(X_1,X_2,X_3,X_5)}}},\\
\boldsymbol{\phi_3} &= \boldsymbol{\mu_{B_{(X_1,X_2,E_{AB},X_5)}} \nu_{C_{(X_1,X_2,E_{AC},X_5)}}},\\
\boldsymbol{\phi_4} &= \boldsymbol{\mu_{B_{(X_1,X_2,X_3,X)}} \nu_{C_{(X_1,X_2,X_3,X)}}};
\end{aligned}$$

und mit (3.27)

$$\boldsymbol{\lambda_A} = \frac{\boldsymbol{V_1 V_2}}{\boldsymbol{V_3 V_4}} \frac{\boldsymbol{\psi_1 \psi_2}}{\boldsymbol{\psi_3 \psi_4}}, \tag{4.10}$$

wobei

$$\begin{aligned}
\boldsymbol{V_1} &= \boldsymbol{T_{BCD} x_{B_{(X_1,X_2,E_{AB},X)}} l_{C_{(X_1,X_2)}} l_{D_{(E_{AD},X)}}},\\
\boldsymbol{V_2} &= \boldsymbol{T_{BCD} x_{B_{(X_1,X_2,X_3,X_5)}} l_{C_{(X_1,X_2)}} l_{D_{(X_3,X_5)}}},\\
\boldsymbol{V_3} &= \boldsymbol{T_{BCD} x_{B_{(X_1,X_2,E_{AB},X_5)}} l_{C_{(X_1,X_2)}} l_{D_{(E_{AD},X_5)}}},\\
\boldsymbol{V_4} &= \boldsymbol{T_{BCD} x_{B_{(X_1,X_2,X_3,X)}} l_{C_{(X_1,X_2)}} l_{D_{(X_3,X)}}},\\
\boldsymbol{\psi_1} &= \boldsymbol{\mu_{B_{(X_1,X_2,E_{AB},X)}}},\\
\boldsymbol{\psi_2} &= \boldsymbol{\mu_{B_{(X_1,X_2,X_3,X_5)}}},\\
\boldsymbol{\psi_3} &= \boldsymbol{\mu_{B_{(X_1,X_2,E_{AB},X_5)}}},\\
\boldsymbol{\psi_4} &= \boldsymbol{\mu_{B_{(X_1,X_2,X_3,X)}}}.
\end{aligned}$$

Sind die Punktkorrespondenzen bekannt und die Epipolar- bzw. Trifokalgeometrie bestimmt, so lassen sich die projektiven Tiefen anhand der Meßdaten leicht berechnen.

4.1.2 Join-Bild

Seien $\boldsymbol{X_j}$, $\boldsymbol{j} \in \{\boldsymbol{1, ..., N}\}$ die $\boldsymbol{N}$ Punkte im Raum und $\boldsymbol{x_{ij}}$, $\boldsymbol{i} \in \{\boldsymbol{1, ..., M}\}$ die Abbildungen dieser Punkte in $\boldsymbol{M}$ Kameras. Die geometrischen Beziehungen zwischen den Punkten im Raum und deren Abbildungen in mehreren Kameras kann man mit Hilfe von bi- und trilinearen Zwangsbedingungen – wie in Kapitel 2 erläutert – beschreiben.

Unter *Join-Bild* wird die Menge der in mehrere Kameras projizierten Punkte (Geraden) einer Szene verstanden. Das Join-Bild wird für die projektive Rekonstruktion einer Szene benutzt. Für diesen Zweck wird das Join-Bild in der Matrixform beschrieben. Dies geschieht, indem man alle Punkte auf eine kohärente Weise in eine Matrix zusammenfügt. Jede Zeile der Join-Bild-Matrix wird aus Elementen der Form

$$\boldsymbol{\lambda}_{ij}\boldsymbol{x}_{ij} = \boldsymbol{P}_i\boldsymbol{X}_j \tag{4.11}$$

gebildet. Die Join-Bild-Matrix $\boldsymbol{\mathcal{J}}$ sieht dann folgendermaßen aus:

$$\boldsymbol{\mathcal{J}}_{3m\times n} = \begin{pmatrix} \boldsymbol{\lambda}_{11}\boldsymbol{x}_{11} & \boldsymbol{\lambda}_{12}\boldsymbol{x}_{12} & \boldsymbol{\lambda}_{13}\boldsymbol{x}_{13} & \cdots & \boldsymbol{\lambda}_{1n}\boldsymbol{x}_{1n} \\ \boldsymbol{\lambda}_{21}\boldsymbol{x}_{21} & \boldsymbol{\lambda}_{22}\boldsymbol{x}_{22} & \boldsymbol{\lambda}_{23}\boldsymbol{x}_{23} & \cdots & \boldsymbol{\lambda}_{2n}\boldsymbol{x}_{2n} \\ \boldsymbol{\lambda}_{31}\boldsymbol{x}_{31} & \boldsymbol{\lambda}_{32}\boldsymbol{x}_{32} & \boldsymbol{\lambda}_{33}\boldsymbol{x}_{33} & \cdots & \boldsymbol{\lambda}_{3n}\boldsymbol{x}_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \boldsymbol{\lambda}_{m1}\boldsymbol{x}_{m1} & \boldsymbol{\lambda}_{m2}\boldsymbol{x}_{m2} & \boldsymbol{\lambda}_{m3}\boldsymbol{x}_{m3} & \cdots & \boldsymbol{\lambda}_{mn}\boldsymbol{x}_{mn} \end{pmatrix}. \tag{4.12}$$

Die $\boldsymbol{\mathcal{J}}$-Matrix hat Rang 4 (es sind die Informationen über die Punkte aus $\mathbb{P}^3 \equiv \mathbb{R}^4$ darin enthalten).

Mit (4.11) läßt sich folgende Beziehung zwischen dem Join-Bild, Kameras und den Punkten im Raum herstellen:

$$\boldsymbol{\mathcal{J}}_{3m\times n} = \begin{pmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \boldsymbol{P}_3 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{P}_m \end{pmatrix}_{3m\times 4} (\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_3 \ldots \boldsymbol{X}_n)_{4\times n}. \tag{4.13}$$

(4.13) ist die Hauptgleichung für die auf der Singulärwertzerlegung basierende projektive Rekonstruktion. Denn sind die Bildkorrespondenzen und die projektiven Tiefen bekannt, so kann man das Join-Bild in die Gestalt- und Bewegungsanteile spalten.

4.1.3 Algorithmus

Somit läßt sich folgender (numerischer) Algorithmus für die projektive Rekonstruktion angeben:

1. Finde in den aufgenommenen Bildern die Punktkorrespondenzen. Anhand dieser Korrespondenzen bestimme die projektiven Tiefen für jeden Bildpunkt in jedem Bild. Diese kann man auf zwei verschiedene Arten finden, wie im Abschnitt 4.1.1 beschrieben ist:

   Mittels Fundamentalmatrix

   (a) Für jedes Bildpaar ermittle die Fundamentalmatrix und die Epipole. Dafür benutze z.B. den Acht-Punkte-Algorithmus [16][3] oder den auf der Basistransformation basierenden Algorithmus [17, 18].

   (b) Setze die projektiven Tiefen von allen Punkten im ersten Bild gleich 1 und berechne die weiteren projektiven Tiefen gemäß der Gleichung (4.4). Die auf diese Weise gefundenen projektiven Tiefen sind bezüglich der ersten Kamera normiert.[4]

   Mittels Invarianten

   (a) Für jedes Bild wähle zwei oder drei weitere Bilder – in der Abhängigkeit davon, ob die Invariante für die projektive Tiefe mittels $\boldsymbol{F}$ oder $\boldsymbol{T}$ berechnet werden soll – und berechne für diese die Fundamentalmatrix bzw. den Trifokaltensor sowie die Epipole. Weiterhin wähle vier Punkte in der allgemeinen Position, die zusammen mit dem optischen Zentrum der Kamera die Basis für die Invariantenberechnung bilden, bzw. ihre Projektionen.

   (b) Berechne die projektiven Tiefen gemäß der Gleichungen (4.9) oder (4.10).

2. Bilde die Join-Bild-Matrix gemäß der Gleichung (4.12).

3. Spalte die $\boldsymbol{\mathcal{J}}$-Matrix mit der Singulärwertzerlegung:

$$\boldsymbol{\mathcal{J}} = \boldsymbol{U diag(\delta_1, ..., \delta_s) V}, \tag{4.14}$$

   wobei $\boldsymbol{s = min(3m, n)}$ und die Singulärwerte $\boldsymbol{\delta_1 \geq ... \geq \delta_s \geq 0}$ sind. Da die linke Seite dieser Gleichung eine Matrix vom Rang 4 ist, sind nur die ersten vier Singulärwerte von Bedeutung, die restlichen Werte gehen gegen 0. Also tragen nur die ersten vier Zeilen (Spalten) der Matrix $\boldsymbol{V(U)}$ zum Produkt bei.

[3]Siehe Anhang D.

[4]Es ist möglich, die nächste Kamera immer bezüglich der vorigen Kamera zu normieren. In diesem Fall spricht man von *serieller* Bearbeitung. Sind dagegen die Tiefen bezüglich nur einer Kamera normiert, so wird von *paralleler* Bearbeitung gesprochen.

4. Bilde die Untermatrizen von $\boldsymbol{U}, \boldsymbol{V}$ und $\boldsymbol{\Delta} = \boldsymbol{diag}(\boldsymbol{\delta_1}, ..., \boldsymbol{\delta_s})$ – $\boldsymbol{U'_{3m\times 4}}, \boldsymbol{V'_{4\times n}}$ und $\boldsymbol{\Delta'} = \boldsymbol{diag}(\boldsymbol{\delta_1}, ..., \boldsymbol{\delta_4})$, und stelle folgende Gleichung auf:

$$\mathcal{J} = \boldsymbol{U'\Delta'V'} = (\boldsymbol{U'\Delta'_1})(\boldsymbol{\Delta'_2 V'}) = \boldsymbol{\hat{P}\hat{X}} = \begin{pmatrix} \boldsymbol{P_1} \\ \boldsymbol{P_2} \\ \vdots \\ \boldsymbol{P_m} \end{pmatrix} \begin{pmatrix} \boldsymbol{X_1} & \boldsymbol{X_2} & \cdots & \boldsymbol{X_n} \end{pmatrix}, \tag{4.15}$$

wobei $\boldsymbol{\Delta'} = \boldsymbol{\Delta'_1 \Delta'_2}$ ist. Da die Rekonstruktion projektiv ist, kann man $\boldsymbol{\Delta'}$ beliebig spalten, z.B. symmetrisch:

$$\boldsymbol{\Delta'_1} = \boldsymbol{\Delta'_2} = (\boldsymbol{\Delta'})^{1/2} = \boldsymbol{diag}(\boldsymbol{\delta_1^{1/2}}, ..., \boldsymbol{\delta_4^{1/2}}).$$

5. Gemäß der Gleichung:

$$\boldsymbol{P_i C_i} = \boldsymbol{0}.$$

finde die optischen Zentren $\boldsymbol{C_i}$'s für alle Kameras.

4.2 Projektive Rekonstruktion mittels Invarianten

In diesem Abschnitt wollen wir die projektive Rekonstruktion mittels Invarianten vorstellen. Einige Autoren haben die Rekonstruktion von Punkten mittels Invarianten schon behandelt. Allerdings wurde diese nur im Raum (und nicht in der Bildebene) ([19]) oder nur für die Punkte (nicht aber für die Kamerapositionen) ([3]) durchgeführt.

Im Folgenden wird gezeigt, wie man die Punkte im Raum anhand der Punktkorrespondenzen in den Bildern mittels Invarianten rekonstruieren kann. Diesem Verfahren zugrunde liegt eine in vorhergegangenen Kapiteln beschriebene Gleichung, die die Koordinaten des zu rekonstruierenden Punktes in Termen der gewählten Raumbasis ausdrückt (Gleichung (3.1)). Weiterhin wird gezeigt, daß sich die Rekonstruktion von optischen Zentren der Kameras (Kamerapositionen) nicht von der Punktrekonstruktion unterscheidet. Anschließend wird ein Algorithmus vorgestellt, der die projektive Rekonstrukion mittels Invarianten vollständig beschreibt.

4.2.1 Rekonstruktion der Gestalt

Betrachte $\boldsymbol{X_j}$, $\boldsymbol{j} \in \{\boldsymbol{1}, ..., \boldsymbol{N}\}$ – die $\boldsymbol{N}$ Punkte im Raum – und $\boldsymbol{x_{ij}}$, $\boldsymbol{i} \in \{\boldsymbol{1}, ..., \boldsymbol{M}\}$ – die Abbildungen dieser Punkte in $\boldsymbol{M}$ Kameras. Wählt man fünf Punkte $\boldsymbol{X_1}$,$\boldsymbol{X_2}$,$\boldsymbol{X_3}$,

$\boldsymbol{X_4}$ und $\boldsymbol{X_5}$ in der allgemeinen Position als die projektive Basis des Raumes, so lassen sich laut (3.1) die Koordinaten jedes weiteren Punktes $\boldsymbol{X_j} = (\mathrm{X_j}, \mathrm{Y_j}, \mathrm{Z_j}, 1)^{\mathrm{T}}$ in Termen dieser Basis schreiben:

$$\mathrm{X_j} = \frac{[234j][1235]}{[2345][123j]}, \quad \mathrm{Y_j} = \frac{[134j][1235]}{[1345][123j]}, \quad \mathrm{Z_j} = \frac{[124j][1235]}{[1245][123j]}. \tag{4.16}$$

Diese Invarianten kann man auch anhand von Punktkorrespondenzen in zwei oder drei Kameras berechnen:

$$\begin{aligned}
\mathrm{X_j} &= \frac{(\boldsymbol{x}^T_{q_{234j}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{234j}})(\boldsymbol{x}^T_{q_{1235}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{1235}})}{(\boldsymbol{x}^T_{q_{2345}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{2345}})(\boldsymbol{x}^T_{q_{123j}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{123j}})} \frac{\mu_{p_{234j}}\mu_{p_{1235}}\nu_{q_{234j}}\nu_{q_{1235}}}{\mu_{p_{2345}}\mu_{p_{123j}}\nu_{q_{2345}}\nu_{q_{123j}}}, \\
\mathrm{Y_j} &= \frac{(\boldsymbol{x}^T_{q_{134j}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{134j}})(\boldsymbol{x}^T_{q_{1235}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{1235}})}{(\boldsymbol{x}^T_{q_{1345}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{1345}})(\boldsymbol{x}^T_{q_{123j}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{123j}})} \frac{\mu_{p_{134j}}\mu_{p_{1235}}\nu_{q_{134j}}\nu_{q_{1235}}}{\mu_{p_{1345}}\mu_{p_{123j}}\nu_{q_{1345}}\nu_{q_{123j}}}, \\
\mathrm{Z_j} &= \frac{(\boldsymbol{x}^T_{q_{124j}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{234j}})(\boldsymbol{x}^T_{q_{1235}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{1235}})}{(\boldsymbol{x}^T_{q_{1245}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{2345}})(\boldsymbol{x}^T_{q_{123j}} \boldsymbol{F}_{pq} \boldsymbol{x}_{p_{123j}})} \frac{\mu_{p_{124j}}\mu_{p_{1235}}\nu_{q_{124j}}\nu_{q_{1235}}}{\mu_{p_{1245}}\mu_{p_{123j}}\nu_{q_{1245}}\nu_{q_{123j}}},
\end{aligned} \tag{4.17}$$

bzw.

$$\begin{aligned}
\mathrm{X_j} &= \frac{(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{234j}}\boldsymbol{l}_{q_{23}}\boldsymbol{l}_{r_{4j}})(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{1235}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{35}})}{(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{2345}}\boldsymbol{l}_{q_{23}}\boldsymbol{l}_{r_{45}})(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{123j}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{3j}})} \frac{\mu_{p_{234j}}\mu_{p_{1235}}}{\mu_{p_{2345}}\mu_{p_{123j}}}, \\
\mathrm{Y_j} &= \frac{(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{134j}}\boldsymbol{l}_{q_{13}}\boldsymbol{l}_{r_{4j}})(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{1235}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{35}})}{(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{1345}}\boldsymbol{l}_{q_{13}}\boldsymbol{l}_{r_{45}})(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{123j}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{3j}})} \frac{\mu_{p_{134j}}\mu_{p_{1235}}}{\mu_{p_{1345}}\mu_{p_{123j}}}, \\
\mathrm{Z_j} &= \frac{(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{124j}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{4j}})(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{1235}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{35}})}{(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{1245}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{45}})(\boldsymbol{T}_{pqr}\boldsymbol{x}_{p_{123j}}\boldsymbol{l}_{q_{12}}\boldsymbol{l}_{r_{3j}})} \frac{\mu_{p_{124j}}\mu_{p_{1235}}}{\mu_{p_{1245}}\mu_{p_{123j}}}.
\end{aligned} \tag{4.18}$$

Die obigen Gleichungen kann man lösen, wenn die Punktkorrespondenzen sowie die Epipolar- bzw. Trifokalgeometrie bekannt sind.

4.2.2 Rekonstruktion der Bewegung

Die projektive Rekonstruktion von optischen Zentren unterscheidet sich auf keine Weise von der Punktrekonstruktion mittels Invarianten – die optischen Zentren sind ganz normale Punkte im Raum. Da die Abbildungen eines optischen Zentrums in den anderen Kameras die Epipole sind, kann man die Koordinaten des optischen Zentrums mittels Invarianten anhand der Bilddaten berechnen.

Sei $\boldsymbol{A_0} = (\mathrm{X_{A_0}}, \mathrm{Y_{A_0}}, \mathrm{Z_{A_0}})^{\mathrm{T}}$ das optische Zentrum der Kamera $\boldsymbol{A}$ und $\boldsymbol{e_{AB}}$,$\boldsymbol{e_{AC}}$,$\boldsymbol{e_{AD}}$ die Epipole von $\boldsymbol{A}$ in den Kameras $\boldsymbol{B}$,$\boldsymbol{C}$ und $\boldsymbol{D}$. Setzt man in (4.17) und (4.18) $\boldsymbol{A_0}$ anstelle von $\boldsymbol{X_j}$ ein, so erhält man die Koordinaten des optischen Zentrums:

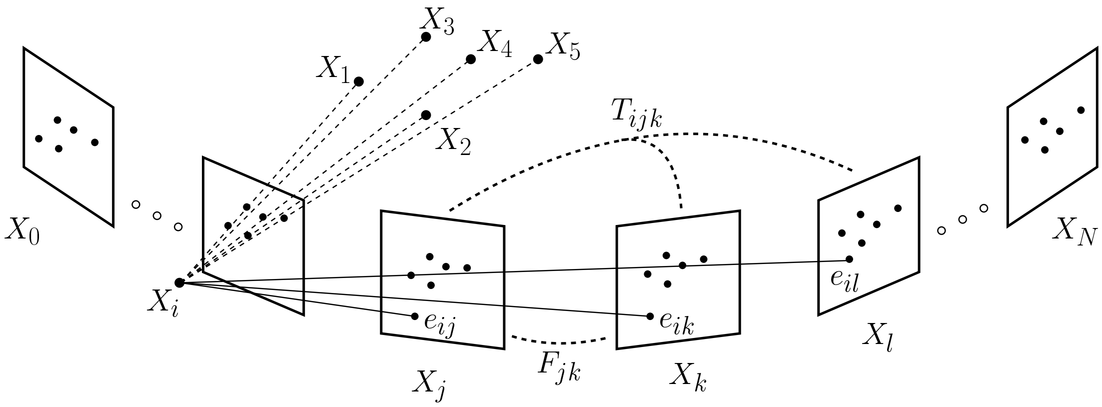


Abbildung 4.3: *Projektive Rekonstruktion mittels Invarianten*

$$\begin{aligned}
\mathbf{X}_{\mathbf{A_0}} &= \frac{(\boldsymbol{x}^T_{C_{234e_{AC}}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{234e_{AB}}})(\boldsymbol{x}^T_{C_{1235}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{1235}})}{(\boldsymbol{x}^T_{C_{2345}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{2345}})(\boldsymbol{x}^T_{C_{123e_{AC}}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{123e_{AB}}})} \cdot \\
&\quad \frac{\boldsymbol{\mu}_{B_{234e_{AB}}}\boldsymbol{\mu}_{B_{1235}}\boldsymbol{\nu}_{C_{234e_{AC}}}\boldsymbol{\nu}_{C_{1235}}}{\boldsymbol{\mu}_{B_{2345}}\boldsymbol{\mu}_{B_{123e_{AB}}}\boldsymbol{\nu}_{C_{2345}}\boldsymbol{\nu}_{C_{123e_{AC}}}}, \\
\mathbf{Y}_{\mathbf{A_0}} &= \frac{(\boldsymbol{x}^T_{C_{134e_{AC}}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{134e_{AB}}})(\boldsymbol{x}^T_{C_{1235}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{1235}})}{(\boldsymbol{x}^T_{C_{1345}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{1345}})(\boldsymbol{x}^T_{C_{123e_{AC}}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{123e_{AB}}})} \cdot \\
&\quad \frac{\boldsymbol{\mu}_{B_{134e_{AB}}}\boldsymbol{\mu}_{B_{1235}}\boldsymbol{\nu}_{C_{134e_{AC}}}\boldsymbol{\nu}_{C_{1235}}}{\boldsymbol{\mu}_{B_{1345}}\boldsymbol{\mu}_{B_{123e_{AB}}}\boldsymbol{\nu}_{C_{1345}}\boldsymbol{\nu}_{C_{123e_{AC}}}}, \\
\mathbf{Z}_{\mathbf{A_0}} &= \frac{(\boldsymbol{x}^T_{C_{124e_{AC}}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{234e_{AB}}})(\boldsymbol{x}^T_{C_{1235}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{1235}})}{(\boldsymbol{x}^T_{C_{1245}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{2345}})(\boldsymbol{x}^T_{C_{123e_{AC}}}\boldsymbol{F}_{BC}\boldsymbol{x}_{B_{123e_{AB}}})} \cdot \\
&\quad \frac{\boldsymbol{\mu}_{B_{124e_{AB}}}\boldsymbol{\mu}_{B_{1235}}\boldsymbol{\nu}_{C_{124e_{AC}}}\boldsymbol{\nu}_{C_{1235}}}{\boldsymbol{\mu}_{B_{1245}}\boldsymbol{\mu}_{B_{123e_{AB}}}\boldsymbol{\nu}_{C_{1245}}\boldsymbol{\nu}_{C_{123e_{AC}}}},
\end{aligned} \tag{4.19}$$

oder

$$\begin{aligned}
\mathbf{X}_{\mathbf{A_0}} &= \frac{(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{234e_{AB}}}\boldsymbol{l}_{C_{23}}\boldsymbol{l}_{D_{4e_{AD}}})(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{1235}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{35}})}{(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{2345}}\boldsymbol{l}_{C_{23}}\boldsymbol{l}_{D_{45}})(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{123e_{AB}}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{3e_{AD}}})} \cdot \\
&\quad \frac{\boldsymbol{\mu}_{B_{234e_{AB}}}\boldsymbol{\mu}_{B_{1235}}}{\boldsymbol{\mu}_{B_{2345}}\boldsymbol{\mu}_{B_{123e_{AB}}}}, \\
\mathbf{Y}_{\mathbf{A_0}} &= \frac{(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{134e_{AB}}}\boldsymbol{l}_{C_{13}}\boldsymbol{l}_{D_{4e_{AD}}})(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{1235}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{35}})}{(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{1345}}\boldsymbol{l}_{C_{13}}\boldsymbol{l}_{D_{45}})(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{123e_{AB}}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{3e_{AD}}})} \cdot \\
&\quad \frac{\boldsymbol{\mu}_{B_{134e_{AB}}}\boldsymbol{\mu}_{B_{1235}}}{\boldsymbol{\mu}_{B_{1345}}\boldsymbol{\mu}_{B_{123e_{AB}}}}, \\
\mathbf{Z}_{\mathbf{A_0}} &= \frac{(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{124e_{AB}}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{4e_{AD}}})(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{1235}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{35}})}{(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{1245}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{45}})(\boldsymbol{T}_{BCD}\boldsymbol{x}_{B_{123e_{AB}}}\boldsymbol{l}_{C_{12}}\boldsymbol{l}_{D_{3e_{AD}}})} \cdot \\
&\quad \frac{\boldsymbol{\mu}_{B_{124e_{AB}}}\boldsymbol{\mu}_{B_{1235}}}{\boldsymbol{\mu}_{B_{1245}}\boldsymbol{\mu}_{B_{123e_{AB}}}}.
\end{aligned} \tag{4.20}$$

Für diese Art der Rekonstruktion benötigt man auf jeden Fall die Epipolargeometrie – die Kenntnisse über die Epipole sind notwendig, unabhängig davon, ob man die Invarianten mittels der Fundamentalmatrix oder des Trifokaltensors berechnen will.

4.2.3 Algorithmus

Somit läßt sich folgendes Verfahren zur projektiven Rekonstruktion mittels Invarianten aufstellen:

1. Finde in den aufgenommenen Bildern die Punktkorrespondenzen. Anhand dieser Korrespondenzen berechne die epipolare bzw. trifokale Geometrie für die Bilder. Wähle fünf Punkte in der allgemeinen Position als die projektive Basis des Raumes.

2. Mit (4.17) oder (4.18) rekonstruiere die Szenenpunkte.

3. Berechne die Kamerapositionen mittels (4.19) bzw. (4.20).

4.3 Erweiterte Rekonstruktion

In den obigen Abschnitten wurde gezeigt, wie man die Rekonstruktion anhand der bekannten Punktkorrespondenzen durchführt. Ist das a priori Wissen über die Szene vorhanden, so lassen sich auch die Punkte rekonstruieren, die in den Bildern nicht zu sehen sind, die sogenannten verdeckten Punkte (*Verdeckungen*). Andererseits kann man die Szene anhand der Analyse der Bildgeometrie um die neuen künstlich gebildeten Punkte erweitern, um so die Umgebung besser beschreiben zu können. Im Folgenden werden einige Fälle von Verdeckungen behandelt, sowie durch einige Beispiele die Methoden der Szenenanalyse demonstriert.

4.3.1 Verdeckte Schnittpunkte

Sei Punkt $\boldsymbol{X}$ im Raum der Schnittpunkt der Geraden $\boldsymbol{L_1}$ und $\boldsymbol{L_2}$. Seien $\boldsymbol{x_A}$,$\boldsymbol{x_B}$,$\boldsymbol{x_C}$ und $\boldsymbol{l_{1_A}}$,$\boldsymbol{l_{2_A}}$,$\boldsymbol{l_{1_B}}$,$\boldsymbol{l_{2_B}}$,$\boldsymbol{l_{1_C}}$,$\boldsymbol{l_{2_C}}$ die Abbildungen dieser Punkte bzw. Geraden in den Kameras $\boldsymbol{A}$,$\boldsymbol{B}$ und $\boldsymbol{C}$.

Ist der Punkt in einem Bild verdeckt, die Geraden aber sichtbar (wie z.B. im Bild $\boldsymbol{\Phi_A}$, Abbildung 4.4), so findet man den verdeckten Punkt als den Schnittpunkt dieser Geraden:

$$\boldsymbol{x_A} = \boldsymbol{l_{1_A}} \bigvee \boldsymbol{l_{2_A}}. \tag{4.21}$$

Sind die Geraden auch verdeckt (Bild $\boldsymbol{\Phi_B}$, Abbildung 4.4), so sind zuerst die verdeckten Geraden gemäß Gleichung (2.35) zu finden:

$$\boldsymbol{l_{1_B}} = \boldsymbol{T_{BAC}} \boldsymbol{l_{1_A}} \boldsymbol{l_{1_C}} \quad \text{bzw.} \quad \boldsymbol{l_{2_B}} = \boldsymbol{T_{BAC}} \boldsymbol{l_{2_A}} \boldsymbol{l_{2_C}}. \tag{4.22}$$

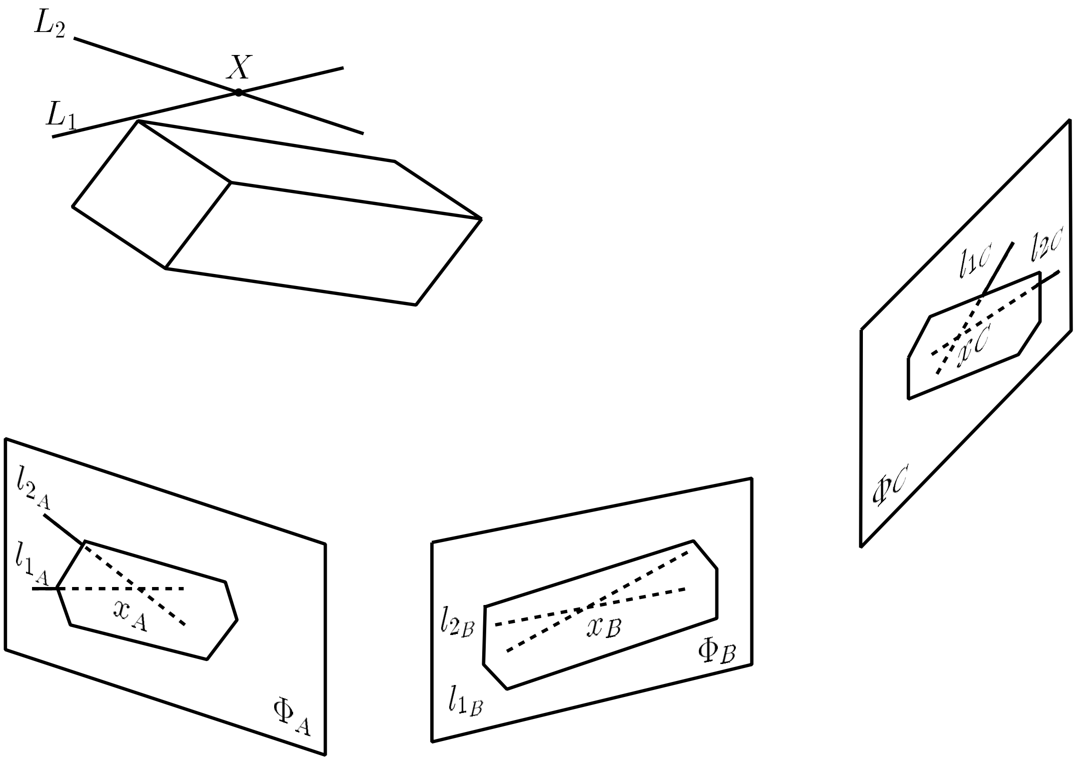


Abbildung 4.4: *Verdeckte Punkte im Bild*

Danach ist der verdeckte Punkt einfach zu bestimmen:

$$\boldsymbol{x_B} = \boldsymbol{l_{1_B}} \bigvee \boldsymbol{l_{2_B}}. \tag{4.23}$$

Sind alle verdeckten Punkte in den Bildern gefunden, kann man die Rekonstruktion wie in den Abschnitten 4.1 und 4.2 beschrieben durchführen.

Die andere Möglichkeit besteht darin, den verdeckten Punkt erst nach der Rekonstruktion zu finden

$$\boldsymbol{X} = \boldsymbol{L_1} \bigvee \boldsymbol{L_2}, \tag{4.24}$$

nachdem die Geraden $\boldsymbol{L_1}$ und $\boldsymbol{L_2}$ (bzw. die Punkte auf diesen Geraden) rekonstruiert wurden.

4.3.2 Verdeckte Eckpunkte

Seien $\boldsymbol{X_1}$,$\boldsymbol{X_2}$,$\boldsymbol{X_3}$,$\boldsymbol{X_4}$,$\boldsymbol{X_5}$,$\boldsymbol{X_6}$ und $\boldsymbol{X}$ wie in der Abbildung 4.5 gegeben. Ist der Punkt $\boldsymbol{X}$ verdeckt, so läßt er sich als Schnittpunkt dreier Ebenen finden:

$$\boldsymbol{X} = \boldsymbol{\Phi_1} \bigvee \boldsymbol{\Phi_2} \bigvee \boldsymbol{\Phi_3}, \tag{4.25}$$

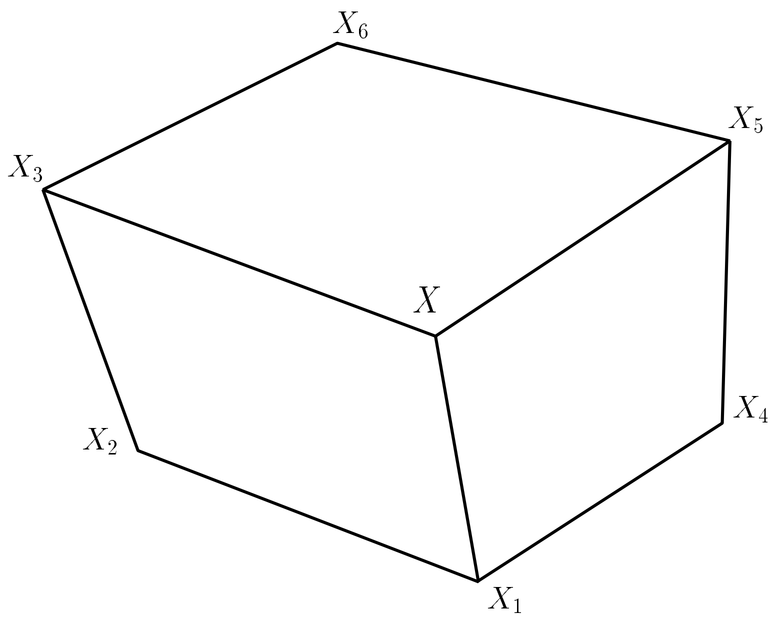


Abbildung 4.5: *Eckpunkte*

wobei

$$\mathbf{\Phi_1} = \boldsymbol{X_1} \bigwedge \boldsymbol{X_2} \bigwedge \boldsymbol{X_3}, \quad \mathbf{\Phi_2} = \boldsymbol{X_1} \bigwedge \boldsymbol{X_4} \bigwedge \boldsymbol{X_5}, \quad \mathbf{\Phi_3} = \boldsymbol{X_3} \bigwedge \boldsymbol{X_5} \bigwedge \boldsymbol{X_6}$$

sind. Sind die Punkte $\boldsymbol{X_1}$,$\boldsymbol{X_2}$,$\boldsymbol{X_3}$,$\boldsymbol{X_4}$,$\boldsymbol{X_5}$ und $\boldsymbol{X_6}$ bekannt, kann man die Gleichung (4.25) direkt benutzen, oder auch $\boldsymbol{X}$ mit Hilfe von Brackets finden:

$$\begin{aligned} [\boldsymbol{X_1 X_2 X_3 X}] &= 0, \\ [\boldsymbol{X_1 X_4 X_5 X}] &= 0, \\ [\boldsymbol{X_3 X_5 X_6 X}] &= 0. \end{aligned} \tag{4.26}$$

Der Punkt $\boldsymbol{X}$ liegt in drei Ebenen $\mathbf{\Phi_1}$,$\mathbf{\Phi_2}$ und $\mathbf{\Phi_3}$, ist also mit entsprechenden Ebenenpunkten koplanar.

4.3.3 Ergänzung der Szene mittels Bildanalyse

Wenn kein a priori Wissen über die Szene vorhanden ist, kann man durch die Analyse der Bilder einige Informationen gewinnen, die es ermöglichen, neue Punkte in die Szene hinzuzufügen. Die wichtigste Information, die man bei der Analyse solcher Art finden kann, ist die Information über die Koplanarität.

Seien $\boldsymbol{L_1}$ und $\boldsymbol{L_2}$ zwei Geraden im Raum und $\boldsymbol{l_{1_A}}$,$\boldsymbol{l_{2_A}}$,$\boldsymbol{l_{1_B}}$,$\boldsymbol{l_{2_B}}$,$\boldsymbol{l_{1_C}}$,$\boldsymbol{l_{2_C}}$ die Abbildungen dieser Geraden in den Kameras $\boldsymbol{A}$,$\boldsymbol{B}$ und $\boldsymbol{C}$. Wählt man auf jeder Gerade je zwei Punkte $\boldsymbol{X_1} \bigwedge \boldsymbol{X_2} = \boldsymbol{L_1}$, $\boldsymbol{X_3} \bigwedge \boldsymbol{X_4} = \boldsymbol{L_2}$, so ist das durch diese vier Punkte gebildete Bracket gleich Null, wenn die Geraden koplanar sind (d.h. wenn sich die Geraden schneiden):

$$[\boldsymbol{X_1X_2X_3X_4}] = 0.$$

In der Bildebene geht man folgenderweise vor, um die Koplanarität nachzuprüfen:

> Man wählt auf $\boldsymbol{l_{1_A}}, \boldsymbol{l_{2_A}}$ je zwei Punkte $\boldsymbol{x_{A_1}} \bigwedge \boldsymbol{x_{A_2}} = \boldsymbol{l_{1_A}}$ und $\boldsymbol{x_{A_3}} \bigwedge \boldsymbol{x_{A_4}} = \boldsymbol{l_{2_A}}$. Die korrespondierenden Punkte in die anderen Kameras findet man mit Hilfe von epipolarer Geometrie – die gesuchten Punkte liegen einerseits auf den Epipolargeraden, andererseits auf den entspechenden Bildgeraden:
>
> $$\begin{aligned} \boldsymbol{x_{B_1}} &= (\boldsymbol{e_{AB}} \bigwedge \boldsymbol{x_{A_1}}) \bigvee \boldsymbol{l_{1_B}}, \\ \boldsymbol{x_{B_2}} &= (\boldsymbol{e_{AB}} \bigwedge \boldsymbol{x_{A_2}}) \bigvee \boldsymbol{l_{1_B}}, \\ \boldsymbol{x_{B_3}} &= (\boldsymbol{e_{AB}} \bigwedge \boldsymbol{x_{A_3}}) \bigvee \boldsymbol{l_{2_B}}, \\ \boldsymbol{x_{B_4}} &= (\boldsymbol{e_{AB}} \bigwedge \boldsymbol{x_{A_4}}) \bigvee \boldsymbol{l_{2_B}}, \\ \boldsymbol{x_{C_1}} &= (\boldsymbol{e_{AC}} \bigwedge \boldsymbol{x_{A_1}}) \bigvee \boldsymbol{l_{1_C}}, \\ \boldsymbol{x_{C_2}} &= (\boldsymbol{e_{AC}} \bigwedge \boldsymbol{x_{A_2}}) \bigvee \boldsymbol{l_{1_C}}, \\ \boldsymbol{x_{C_3}} &= (\boldsymbol{e_{AC}} \bigwedge \boldsymbol{x_{A_3}}) \bigvee \boldsymbol{l_{2_C}}, \\ \boldsymbol{x_{C_4}} &= (\boldsymbol{e_{AC}} \bigwedge \boldsymbol{x_{A_4}}) \bigvee \boldsymbol{l_{2_C}}. \end{aligned} \tag{4.27}$$
>
> Mit Hilfe von (3.13) und (3.23) überprüft man die Punkte auf Koplanarität:
>
> $$\boldsymbol{x^T_{B_{1234}} F_{AB} x_{A_{1234}} = 0} \tag{4.28}$$
>
> und
>
> $$\boldsymbol{T_{ABC} x_{A_{1234}} l_{B_{12}} l_{C_{34}} = 0}, \tag{4.29}$$
>
> wobei die Korrekturkoeffizenten keine Rolle spielen, da mit Null verglichen wird. Um sicher zu sein, daß die Geraden koplanar sind und keine ausgearteten Bilder benutzt wurden, sollte man den Test in mehreren Kameras und möglichst mit Tensor durchführen.

Sind die Koplanaritäten sichergestellt, so geht man weiter vor wie in den Abschnitten 4.3.1 und 4.3.2.

4.4 Euklidische Rekonstruktion

Für die praktische Anwendung ist die euklidische Rekonstruktion von großer Bedeutung. Deshalb wollen wir in diesem Abschnitt die Möglichkeiten vorstellen, eine projektive Rekonstruktion in eine euklidische zu transformieren. Dieses geschieht, indem man eine projektive Transformation findet, die die rekonstruierten Punkte in eine euklidische Basis des Raumes überführt. Ein Weg, eine solche Transformation zu finden, ist, die euklidischen Koordinaten von mindestens fünf Punkten in der allgemeinen Position zu kennen. Dieses Wissen kann a priori Wissen sein, oder durch die Auswertung der Bilddaten gewonnen werden.

Seien $\tilde{\boldsymbol{X}}_{\boldsymbol{i}} = (\tilde{\mathbf{X}}_\mathbf{i}, \tilde{\mathbf{Y}}_\mathbf{i}, \tilde{\mathbf{Z}}_\mathbf{i}, \tilde{\mathbf{W}}_\mathbf{i})^\mathbf{T}$ und $\boldsymbol{X}_{\boldsymbol{i}} = (\mathbf{X}_\mathbf{i}, \mathbf{Y}_\mathbf{i}, \mathbf{Z}_\mathbf{i}, \mathbf{W}_\mathbf{i})^\mathbf{T}$ die Repräsentationen eines Punktes in einer projektiven bzw. euklidischen Basis. Bezeichnet man mit

$$\boldsymbol{T}_{\boldsymbol{p2e}} = \begin{pmatrix} \boldsymbol{t}_{\mathbf{11}} & \boldsymbol{t}_{\mathbf{12}} & \boldsymbol{t}_{\mathbf{13}} & \boldsymbol{t}_{\mathbf{14}} \\ \boldsymbol{t}_{\mathbf{21}} & \boldsymbol{t}_{\mathbf{22}} & \boldsymbol{t}_{\mathbf{23}} & \boldsymbol{t}_{\mathbf{24}} \\ \boldsymbol{t}_{\mathbf{31}} & \boldsymbol{t}_{\mathbf{32}} & \boldsymbol{t}_{\mathbf{33}} & \boldsymbol{t}_{\mathbf{34}} \\ \boldsymbol{t}_{\mathbf{41}} & \boldsymbol{t}_{\mathbf{42}} & \boldsymbol{t}_{\mathbf{43}} & \boldsymbol{t}_{\mathbf{44}} \end{pmatrix}$$

die Transformation zwischen diesen Basen, so gilt:

$$\mathbf{S_i}\boldsymbol{X}_{\boldsymbol{i}} = \boldsymbol{T}_{\boldsymbol{p2e}}\tilde{\boldsymbol{X}}_{\boldsymbol{i}},$$

wobei $\mathbf{S_i}$'s Skalarfaktoren sind. Somit ergibt sich:

$$\begin{aligned} \mathbf{S_i X_i} &= \boldsymbol{t}_{\mathbf{11}}\tilde{\mathbf{X}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{12}}\tilde{\mathbf{Y}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{13}}\tilde{\mathbf{Z}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{14}}\tilde{\mathbf{W}}_\mathbf{i}, \\ \mathbf{S_i Y_i} &= \boldsymbol{t}_{\mathbf{21}}\tilde{\mathbf{X}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{22}}\tilde{\mathbf{Y}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{23}}\tilde{\mathbf{Z}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{24}}\tilde{\mathbf{W}}_\mathbf{i}, \\ \mathbf{S_i Z_i} &= \boldsymbol{t}_{\mathbf{31}}\tilde{\mathbf{X}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{32}}\tilde{\mathbf{Y}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{33}}\tilde{\mathbf{Z}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{34}}\tilde{\mathbf{W}}_\mathbf{i}, \\ \mathbf{S_i W_i} &= \boldsymbol{t}_{\mathbf{41}}\tilde{\mathbf{X}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{42}}\tilde{\mathbf{Y}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{43}}\tilde{\mathbf{Z}}_\mathbf{i} + \boldsymbol{t}_{\mathbf{44}}\tilde{\mathbf{W}}_\mathbf{i}. \end{aligned} \tag{4.30}$$

Also liefert jeder Punkt vier Gleichungen, die bezüglich der Koeffizienten der Transformationsmatrix linear sind, und eine Unbekannte. Da die projektive Transformation bis auf einen Skalarfaktor definiert ist, braucht man mindestens fünf Punkte, um diese zu finden.

Kapitel 5

Visuell basierte Robotermanipulationen

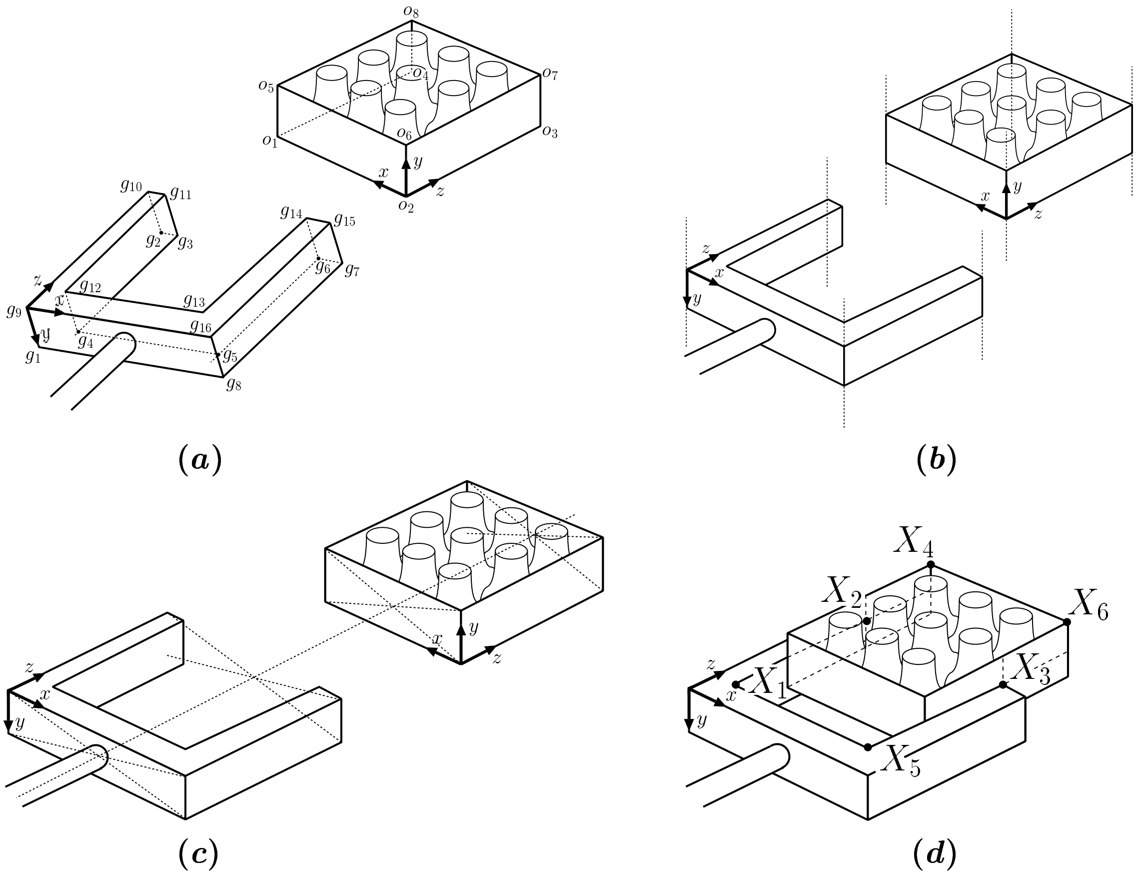


Abbildung 5.1: *Greifen eines Objektes:* **(*a*)** *Allgemeine Position* **(*b*)** *Parallele Orientierung* **(*c*)** *Zentrierung* **(*d*)** *Greifen*

Der Zweck einer Rekonstruktion besteht darin, die Umgebung kennenzulernen (zu rekonstruieren), um mit dieser agieren zu können. In diesem Kapitel wollen wir die Möglichkeiten der geometrischen Algebra und der Invariantztheorie einsetzen, um zu demonstrieren, wie man anhand der visuellen Daten die Robotermanipulationen durchführen kann. Dabei wird das a priori Wissen über die geometrische Struktur des Robotergreifers und des zu greifenden Objektes, nicht jedoch über ihre Positionierung zueinander, vorausgesetzt (siehe Abbildung 5.1). Weiterhin beschränkt sich das Verfahren, das wir hier vorstellen möchten, auf die Szenenanalyse. D.h. die Position von Objekt und Greifer zueinander wird geschätzt, es werden jedoch keine Vorschläge zur gezielten Steuerung des Greifers gemacht.

Das Greifen eines Objektes kann in mehrere Schritte unterteilt werden:

1. Parallele Orientierung – alle Flächen des Greifers und des Objektes werden parallel zueinander ausgerichtet

2. Zentrierung – Der Greifer wird gegenüber dem Objekt positioniert

3. Greifen – die inneren Flächen des Greifers und das Objekt berühren sich möglichst zentral

Dabei sollen für jede Bedingung die vorherigen Bedingungen erhalten bleiben.

5.1 Parallele Orientierung

Betrachte Abbildung 5.1. Seien $\{\boldsymbol{o}_{\boldsymbol{A}_i}\}$,$\{\boldsymbol{g}_{\boldsymbol{A}_i}\}$, $\{\boldsymbol{o}_{\boldsymbol{B}_i}\}$,$\{\boldsymbol{g}_{\boldsymbol{B}_i}\}$ und $\{\boldsymbol{o}_{\boldsymbol{C}_i}\}$,$\{\boldsymbol{g}_{\boldsymbol{C}_i}\}$ die Abbildungen der Szene in den Kameras $\boldsymbol{A}$,$\boldsymbol{B}$ und $\boldsymbol{C}$. Definiere die Fluchtpunkte $\boldsymbol{v_x}$,$\boldsymbol{v_y}$,$\boldsymbol{v_z}$ (die Punkte im Unendlichen, wo sich die parallelen Geraden schneiden) für die orthogonalen Kanten des Objektes:

$$\begin{aligned}
\boldsymbol{v_x} &= (\boldsymbol{o_1} \bigwedge \boldsymbol{o_2}) \bigvee (\boldsymbol{o_5} \bigwedge \boldsymbol{o_6}), \\
\boldsymbol{v_y} &= (\boldsymbol{o_1} \bigwedge \boldsymbol{o_5}) \bigvee (\boldsymbol{o_2} \bigwedge \boldsymbol{o_6}), \\
\boldsymbol{v_x} &= (\boldsymbol{o_1} \bigwedge \boldsymbol{o_4}) \bigvee (\boldsymbol{o_2} \bigwedge \boldsymbol{o_3}),
\end{aligned} \tag{5.1}$$

sowie die entsprechenden Punkte in der Bildebene

$$\begin{aligned}
\boldsymbol{v_{j_x}} &= (\boldsymbol{o_{j_1}} \bigwedge \boldsymbol{o_{j_2}}) \bigvee (\boldsymbol{o_{j_5}} \bigwedge \boldsymbol{o_{j_6}}), \\
\boldsymbol{v_{j_y}} &= (\boldsymbol{o_{j_1}} \bigwedge \boldsymbol{o_{j_5}}) \bigvee (\boldsymbol{o_{j_2}} \bigwedge \boldsymbol{o_{j_6}}), \\
\boldsymbol{v_{j_z}} &= (\boldsymbol{o_{j_1}} \bigwedge \boldsymbol{o_{j_4}}) \bigvee (\boldsymbol{o_{j_2}} \bigwedge \boldsymbol{o_{j_3}}),
\end{aligned} \tag{5.2}$$

für $\boldsymbol{j} \in \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}$.

Die Parallelitätsbedingung läßt sich dann auf zwei verschiedene Weisen überprüfen:

- Die orthogonalen Flächen des Greifers und des Objektes sind parallel, was in den folgenden Gleichungen beschrieben ist:

$$
\begin{aligned}
[\boldsymbol{g}_1\boldsymbol{g}_8\boldsymbol{o}_1\boldsymbol{o}_2] &= \mathbf{0}, \\
[\boldsymbol{g}_{15}\boldsymbol{g}_{16}\boldsymbol{o}_5\boldsymbol{o}_8] &= \mathbf{0}, \\
[\boldsymbol{g}_{12}\boldsymbol{g}_{13}\boldsymbol{o}_3\boldsymbol{o}_4] &= \mathbf{0}
\end{aligned} \tag{5.3}
$$

- Die orthogonalen Kanten des Greifers sind zu den Kanten des Objektes parallel:

$$
\begin{aligned}
(\boldsymbol{g}_1 \bigwedge \boldsymbol{g}_8) \bigwedge \boldsymbol{v}_x &= \mathbf{0}, \\
(\boldsymbol{g}_1 \bigwedge \boldsymbol{g}_9) \bigwedge \boldsymbol{v}_y &= \mathbf{0}, \\
(\boldsymbol{g}_1 \bigwedge \boldsymbol{g}_2) \bigwedge \boldsymbol{v}_z &= \mathbf{0}
\end{aligned} \tag{5.4}
$$

Diese Bedingungen kann man auch in Termen der Bildkoordinaten aufschreiben. Die Bedingungen (5.3) werden zu

$$
\begin{aligned}
\boldsymbol{x}^T_{j_{g_1 g_8 o_1 o_2}} \boldsymbol{F}_{ij} \boldsymbol{x}_{i_{g_1 g_8 o_1 o_2}} &= \mathbf{0}, \\
\boldsymbol{x}^T_{j_{g_{15} g_{16} o_5 o_8}} \boldsymbol{F}_{ij} \boldsymbol{x}_{i_{g_{15} g_{16} o_5 o_8}} &= \mathbf{0}, \\
\boldsymbol{x}^T_{j_{g_{12} g_{13} o_3 o_4}} \boldsymbol{F}_{ij} \boldsymbol{x}_{i_{g_{12} g_{13} o_3 o_4}} &= \mathbf{0},
\end{aligned} \tag{5.5}
$$

oder

$$
\begin{aligned}
\boldsymbol{T}_{ijk} \boldsymbol{x}_{i_{g_1 g_8 o_1 o_2}} \boldsymbol{l}_{j_{g_1 g_8}} \boldsymbol{l}_{k_{o_1 o_2}} &= \mathbf{0}, \\
\boldsymbol{T}_{ijk} \boldsymbol{x}_{i_{g_{15} g_{16} o_5 o_8}} \boldsymbol{l}_{j_{g_{15} g_{16}}} \boldsymbol{l}_{k_{o_5 o_8}} &= \mathbf{0}, \\
\boldsymbol{T}_{ijk} \boldsymbol{x}_{i_{g_{12} g_{13} o_3 o_4}} \boldsymbol{l}_{j_{g_{12} g_{13}}} \boldsymbol{l}_{k_{o_3 o_4}} &= \mathbf{0}.
\end{aligned} \tag{5.6}
$$

Und die Bedingungen (5.4) werden zu:

$$
\begin{aligned}
[\boldsymbol{g}_{i_1} \boldsymbol{g}_{i_8} \boldsymbol{v}_{i_x}] &= \mathbf{0}, \\
[\boldsymbol{g}_{i_1} \boldsymbol{g}_{i_9} \boldsymbol{v}_{i_y}] &= \mathbf{0}, \\
[\boldsymbol{g}_{i_1} \boldsymbol{g}_{i_2} \boldsymbol{v}_{i_z}] &= \mathbf{0}.
\end{aligned} \tag{5.7}
$$

Für die letzten Gleichungen braucht man keine Kenntnisse über die Fundamentalmatrix oder den Trifokaltensor – die Berechnungen können mit den Punkten in lediglich einer Kamera durchgeführt werden.

5.2 Zentrierung

Die Zentrierungsbedingung kann man auf folgende Weise beschreiben:

Die Z-Achse des Greifers und des Objektes sind kollinear.

Definiert man die 'Zentrierungspunkte' wie folgt

$$\begin{aligned} \boldsymbol{c}_o &= (\boldsymbol{o}_1 \bigwedge \boldsymbol{o}_6) \bigvee (\boldsymbol{o}_2 \bigwedge \boldsymbol{o}_5), \\ \boldsymbol{c}_g &= (\boldsymbol{g}_1 \bigwedge \boldsymbol{g}_{16}) \bigvee (\boldsymbol{g}_8 \bigwedge \boldsymbol{g}_9), \end{aligned}$$

dann sieht die obige Bedingung so aus:

$$\boldsymbol{c}_o \bigwedge \boldsymbol{c}_g \bigwedge \boldsymbol{v}_z = \mathbf{0}. \tag{5.8}$$

In der Bildebene kann man (5.8) in Termen des Brackets ausdrücken:

$$[\boldsymbol{c}_{i_o} \boldsymbol{c}_{i_g} \boldsymbol{v}_{i_z}] = \mathbf{0}. \tag{5.9}$$

Das obige Bracket kann mit Punkten in einer Kamera berechnet werden.

5.3 Greifen

Das Greifen wird als die Berührung der Flächen des Greifers und des Objektes definiert. Dies kann als eine Koplanaritätsbedingung aufgeschrieben werden, z.B.:

$$[\boldsymbol{c}_o \boldsymbol{c}_g \boldsymbol{o}_1 \boldsymbol{o}_2] = \mathbf{0}. \tag{5.10}$$

In Termen des Tensors und der Fundamentalmatrix wird (5.10) folgendermaßen ausgedrückt:

$$\begin{aligned} \boldsymbol{x}^T_{j_{c_o c_g o_1 o_2}} \boldsymbol{F}_{ij} \boldsymbol{x}_{i_{c_o c_g o_1 o_2}} &= \mathbf{0}, \\ \boldsymbol{T}_{ijk} \boldsymbol{x}_{i_{c_o c_g o_1 o_2}} \boldsymbol{l}_{j_{c_o c_g}} \boldsymbol{l}_{k_{o_1 o_2}} &= \mathbf{0}. \end{aligned} \tag{5.11}$$

Diese Bedingung kann überprüft werden, wenn die Epipolar- bzw. Trifokalgeometrie bekannt ist.

Kapitel 6

Experimente und Ergebnisse

Die theoretischen Grundlagen der Invarianz- sowie der SVD-Rekonstruktion wurden von uns in den vorangegangenen Kapiteln vollständig behandelt. Nun wollen wir diese Konzepte mit experimentellen Ergebnissen illustrieren. Dabei werden zunächst künstlich erzeugte rauschfreie Bilder benutzt. Weiterhin wird das Verhalten von verschiedenen Rekonstruktionsverfahren unter dem Einfluß von Rauschen analysiert. Anschließend werden echte Daten für die Rekonstruktion benutzt.

Am Anfang wollen wir die Stabilität von Invarianten analysieren. Danach wird die Rolle der projektiven Tiefe anhand der Experimente erklärt. Zum Schluss werden die Möglichkeiten der SVD- und Invariantenrekonstruktion offengelegt.

6.1 Invarianten

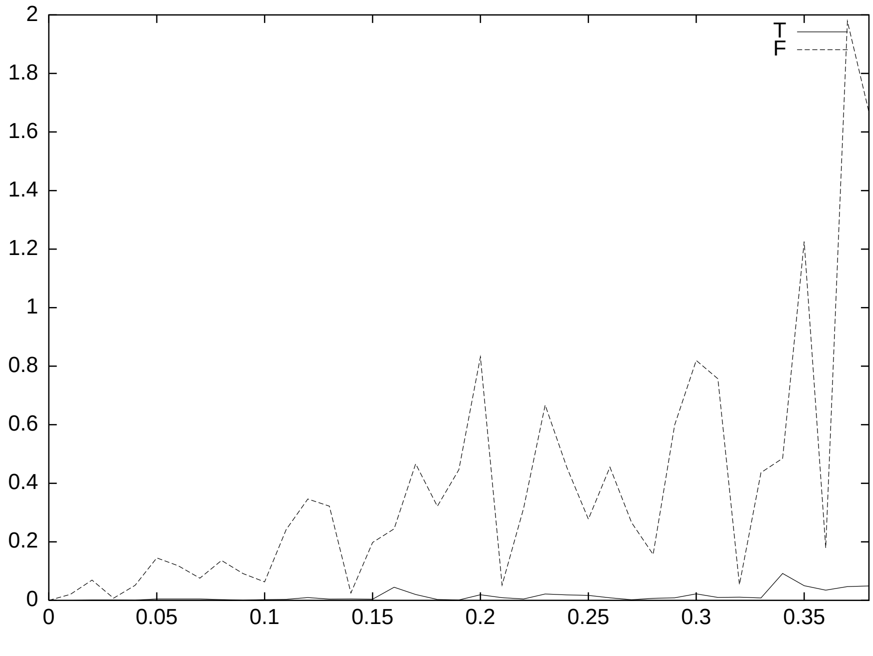
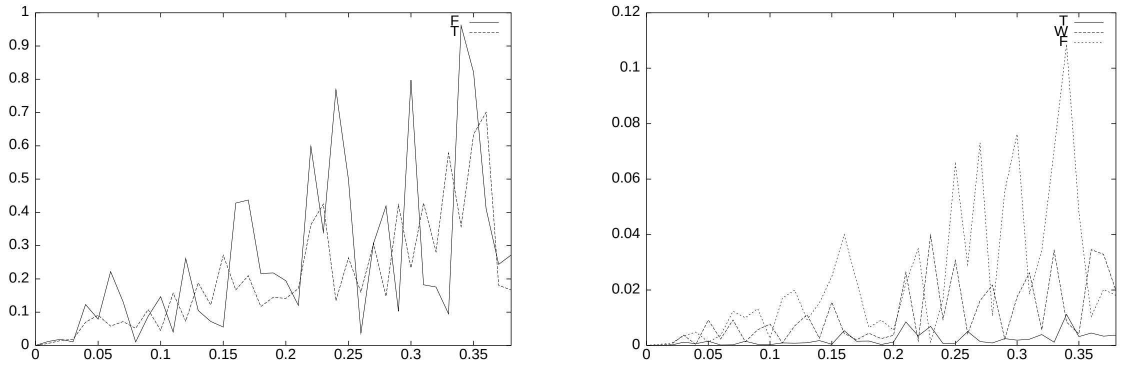


Abbildung 6.1: *Koplanaritätstest mit* $\boldsymbol{F}$ *und* $\boldsymbol{T}$ *– der absolute Fehler (Ordinate) in Abhängigkeit zu weissem Rauschen (Abzisse)*

Eine der für uns wichtigsten Invarianten ist die Koplanaritätsinvariante – ein Bracket, das mit Null verglichen wird. Diese Invariante ist aus mehreren Gründen interessant. Sie ist für die Wahl der Punkte in der allgemeinen Position sowie für die erweiterte Rekonstruktion von großer Bedeutung. Deshalb wollen wir das Verhalten der Koplanaritätsinvariante unter der Wirkung von Rauschen untersuchen.

Die Abbildung 6.1 zeigt die absoluten Werte des Brackets von vier koplanaren Punkten, berechnet mit Hilfe der Fundamentalmatrix und des Trifokaltensors, unter dem Einfluß zunehmenden Rauschens. Die Abzisse zeigt die Werte des Rauschens in Prozent bezüglich der Bilddimension an. Der Test wurde mit *MAPLE V* durchgeführt. Für das Experiment wurde eine Wolke aus zufälligen Punkten generiert und in mehrere virtuelle Kameras projiziert. Vier Punkte aus dieser Wolke wurden aber koplanar gewählt. Die Fundamentalmatrizen für die verschiedenen Bildpaare sowie die Trifokaltensoren wurden rauschfrei berechnet. Für die Berechnung des Brackets wurden die Gleichungen (3.13) bzw. (3.23) benutzt. Die Korrekturkoeffizienten wurden jedoch nicht berücksichtigt, da sie keine Rolle bei dem Vergleich mit Null spielen. Das Rauschen wurde in den Bildern zunehmend erhöht. Für die Berechnung der Invariante wurden immer rauschfreie $\boldsymbol{F}$'s und $\boldsymbol{T}$'s benutzt.



$(\boldsymbol{a})$ $(\boldsymbol{b})$

Abbildung 6.2: *Invarianten unter der Einwirkung von Rauschen:* $(\boldsymbol{a})$ $\boldsymbol{Inv_{2cop}}$ $(\boldsymbol{b})$ $\boldsymbol{Inv_{cop1}}$

In Abbildung 6.2 sind die Ergebnisse der Rauscheinwirkung auf die Invarianten $\boldsymbol{Inv_{cop1}}$ (Gleichung (3.8)) und $\boldsymbol{Inv_{2cop}}$ (Gleichung (3.11)) dargestellt. Die Abzisse zeigt die Werte des Rauschens in Prozent bezüglich der Bilddimension an. Die Ordinate beschreibt den relativen Fehler. Für alle Berechnungen wurden rauschfreie Fundamentalmatrizen bzw. Trifokaltensoren benutzt.

Es hat sich herausgestellt, daß in allen Fällen die Abschätzung der Invarianten mit dem Trifokaltensor besser ist als mit den Fundamentalmatrizen. Allerdings sind die erzielten Ergebnisse nicht sehr gut ausgefallen – bei dem 0.4-prozentigen Rauschen (für das 500×500-Bild sind das 2 Pixel) beträgt die Abweichung von dem echten Wert über hundert Prozent. Das für den Koplanaritätstest verwendete Bracket ist dagegen auch für die echten Daten zu gebrauchen. Die Qualität der ermittelten Fundamentalmatrizen und Tensoren spielt dabei eine sehr große Rolle.

6.2 SVD-Rekonstruktion

Wie in Abschnitt 4 beschrieben wurde, besteht die projektive Rekonstruktion im wesentlichen aus zwei Schritten – der Bestimmung der projektiven Tiefe und der Spaltung des Join-Bildes in Bewegungs- und Gestaltanteil. Im weiteren wollen wir demonstrieren, welche Rolle die einzelnen Schritte für die gesamte Rekonstruktion spielen.

6.2.1 Rolle der projektiven Tiefe für die Rekonstruktion

Die projektive Tiefe beschreibt die Beziehung zwischen dem Abstand eines Punktes zur Kamera und der Brennweite dieser Kamera. Folglich spielt die projektive Tiefe für die Rekonstruktion eine Rolle, wenn sie für die verschiedenen Szenenpunkte stark variiert. Sind also alle Punkte der Szene kompakt plaziert und weit von der Kamera entfernt, so sind die projektiven Tiefen dieser Punkte fast gleich. Sind dagegen die Punkte der Kamera nah oder im Raum stark verstreut, so ist die Tiefenvariation auch groß und für die Rekonstruktion von Bedeutung.

Dies wird in der Abbildung 6.3 illustriert. In diesem Experiment wurde das Modell in die neun virtuellen Kameras projiziert. Alle Kameras haben gleiche intrinsische Parameter und sind für alle Aufnahmen fest plaziert. Die Brennweite beträgt fünf Einheiten. Das Variieren der Distanz erfolgt nur über die Verschiebung des Objektes.

Wie man sieht, verliert die projektive Tiefe bei der zunehmenden Entfernung immer mehr an Bedeutung für die Rekonstruktion. Man kann diese Tatsache für die Schätzung der Entfernung nutzen. Allerdings benötigt dies doppelten Aufwand – die Rekonstruktion wird einmal mit und einmal ohne die Tiefen durchgeführt. Sind die Ergebnisse ähnlich, so ist das Objekt weit entfernt. Unterscheiden sich die rekonstruierten Objekte stark voneinander, so ist das Objekt ganz nah.

6.2.2 Eigenschaften der SVD-Rekonstruktion

Die SVD-Rekonstruktion ist eine batch-Rekonstruktion, d.h. alle Punkte samt aller Kameras werden gleichzeitig verarbeitet. Solche batch-Verfahren haben die Eigenschaft, daß alle Fehler (Rauschen) gemittelt werden. Dies bedeutet einerseits, daß die 'schlechteren' Punktkorrespondenzen durch die 'besseren' ausgeglichen werden. Andererseits können die 'guten' Korrespondenzen durch ein paar 'sehr schlechte' Korrespondenzen zu stark verrauscht werden.

Die Abbildung 6.4 demonstriert die Ergebnisse eines Experimentes mit den verrauschten Daten. In diesem Experiment wurden die Auswirkungen des Rauschens auf die Rekonstruktionsergebnisse untersucht. Das Experiment lief folgendermaßen ab:

- Ein künstliches Objekt, bestehend aus 109 Punkten, wurde in die virtuellen Kameras projiziert. Mit den auf diese Weise gewonnenen Daten wurde die SVD-Rekonstruktion durchgeführt (Abbildung 6.4$\boldsymbol{a}$)

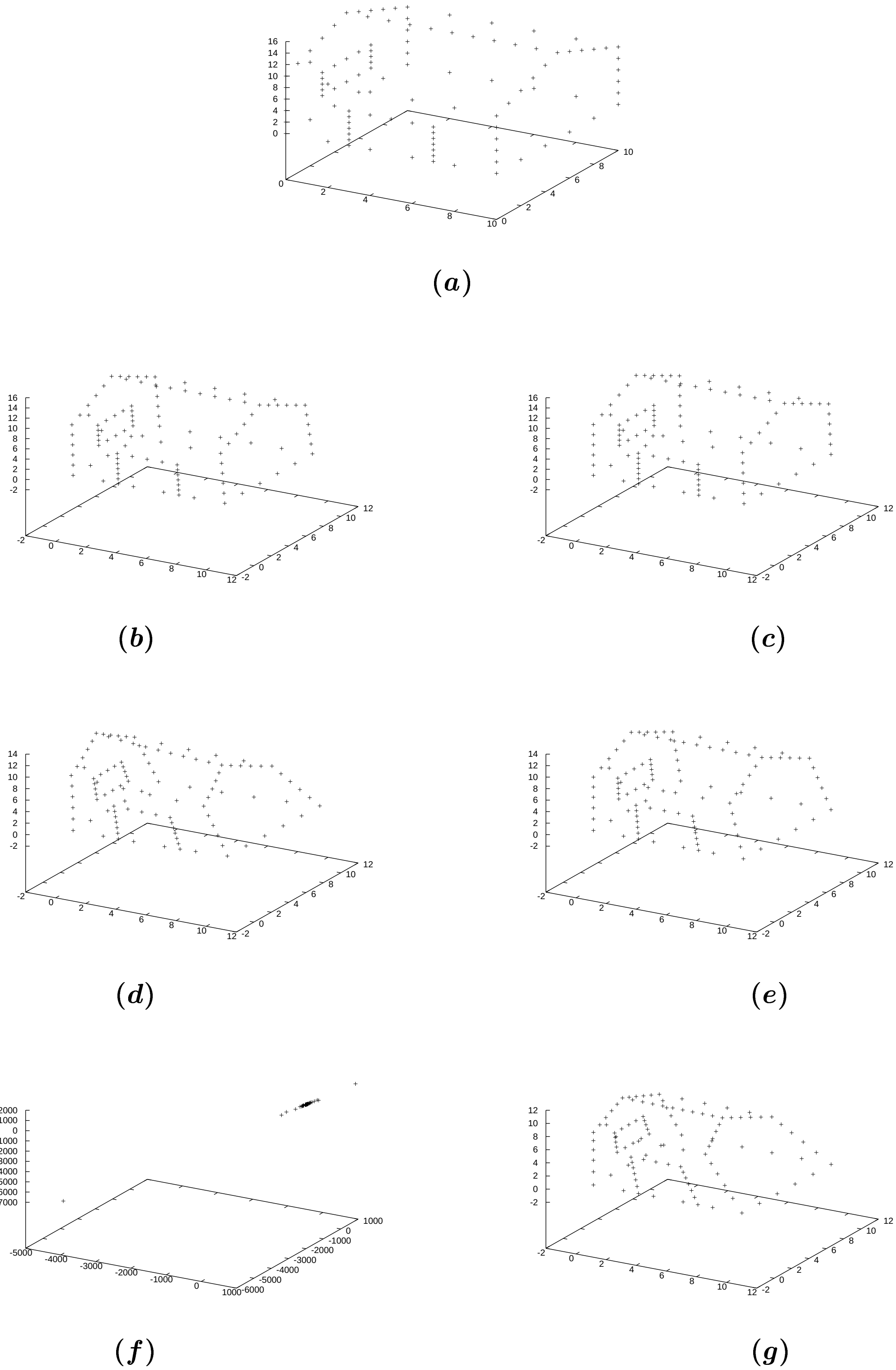
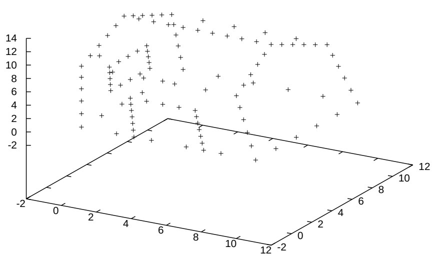
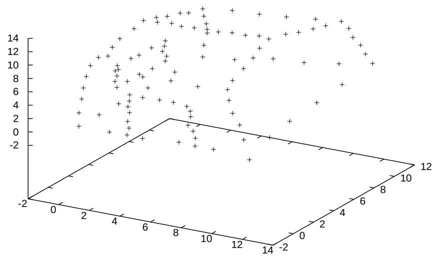
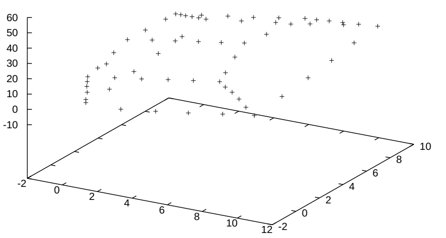


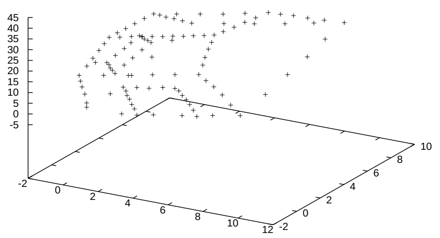
Abbildung 6.3: *Rolle der projektiven Tiefe:* ***(a)*** *Raummodell* ***(b)*** *Distanz* $1000-1500$ *– Rekonstruktion ohne projektive Tiefen* ***(c)*** *Distanz* $1000-1500$ *– Rekonstruktion mit projektiven Tiefen* ***(d)*** *Distanz* $100-200$ *– Rekonstruktion ohne projektive Tiefen* ***(e)*** *Distanz* $100-200$ *– Rekonstruktion mit projektiven Tiefen* ***(f)*** *Distanz* $30-35$ *– Rekonstruktion ohne projektive Tiefen* ***(g)*** *Distanz* $30-35$ *– Rekonstruktion mit projektiven Tiefen*





(***a***) (***b***)





(***c***) (***d***)

Abbildung 6.4: *Rolle des Rauschens:* (***a***) *Rauscharm* (***b***) *Erhöhtes Rauschen (0.2 Prozent)* (***c***) *Erhöhtes Rauschen (0.2 Prozent) mit einem stark verrauschten Punkt (8 Prozent)* (***d***) *Erhöhtes Rauschen (0.2 Prozent) mit einem stark verrauschten Punkt (8 Prozent) und zusätzlichen Punkten (0.2 Prozent verrauscht)*

- Die Bilddaten wurden verrauscht (0.2 Prozent bezüglich der Bilddimension). Das Ergebnis der Rekonstruktion von 109 Punkten ist in der Abbildung 6.4***b*** zu sehen

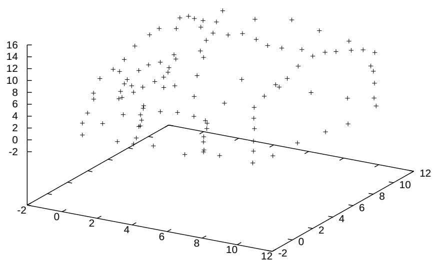
- Eine Punktkorrespondenz wurde zufällig ausgewählt und mit erhöhtem Rauschen (8 Prozent bezüglich der Bilddimension) belegt. Dabei wurden für die Rekonstruktion 65 Punktkorrespondenzen benutzt (Abbildung 6.4***c***)

- Anschließend wurden alle 109 Punktkorrespondenzen in die Rekonstruktion miteinbezogen (siehe Abbildung 6.4***d***)

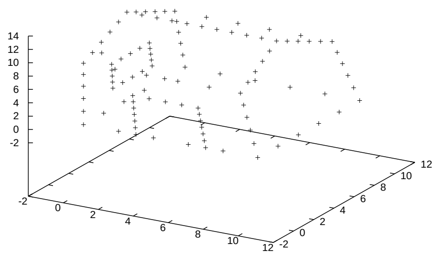
Wie man den Ergebnissen entnehmen kann, wirkt das Einfügen von 'guten' Bildkorrespondenzen positiv auf die gesamte Rekonstruktion. Man sollte also für die Rekonstruktion möglichst viele Punkte benutzen. Und diese Punkte sollten möglichst viele 'gute' Korrespondenzen bieten. Die Überprüfung einer Korrespondenz auf die Eignung kann man mittels Epipolar- oder Trifokalgeometrie durchführen.

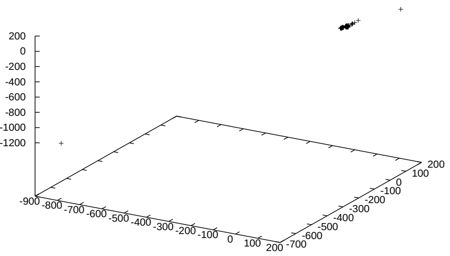
Gibt es zu wenig Punkte mit 'guter' Korrespondenz, so kann man die neuen Korrespondenzen mittels geometrischer Algebra oder der Invarianztheorie aus vorhandenen 'guten' Korrespondenzen selbst erzeugen.

Die Anzahl der Bilder spielt auch eine Rolle – je mehr Bilder für die Rekonstruktion miteinbezogen werden, desto größer ist die Wahrscheinlichkeit, daß die guten Korrespondenzen gefunden werden können. Das liegt daran, daß durch die verschiedenen Ansichten die aufgenommenen Punkte in einigen Aufnahmen besser zu finden sind. Die Abbildung 6.5 demonstriert die Ergebnisse der Rekonstruktion mit verschiedener Anzahl von Kameras.



(a)



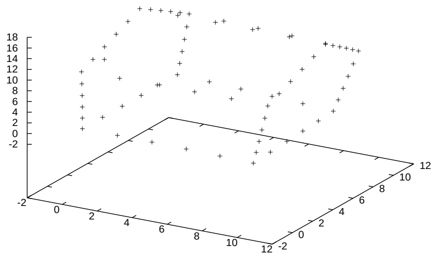


(b) ***(c)***

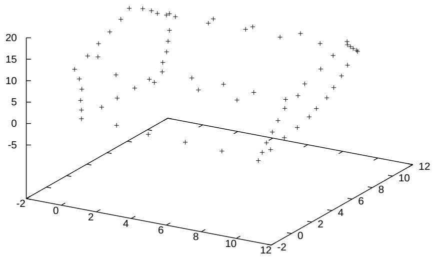
Abbildung 6.5: *Rolle der Kameras:* ***(a)*** *3 Kameras mit dem Rauschniveau 0.2 Prozent* ***(b)*** *6 Kameras – 3 Kameras wie in* ***(a)*** *und 3 mit dem Rauschniveau 0.1 Prozent* ***(c)*** *9 Kameras – 6 Kameras wie in* ***(b)*** *und 3 mit dem Rauschniveau 1 Prozent*

In diesem Experiment wurde ein Objekt mit neun Kameras aufgenommen. Die ersten drei Bilder wurden mit 0.2-prozentigem Rauschen belegt. Die nächsten drei Bilder wurden mit 0.1 und die letzen drei mit 1 Prozent verrauscht. Die Ergebnisse zeigen, daß sich das Hinzufügen von rauscharmen Bildern vorteilhaft auf die gesamte Rekonstruktion auswirkt. Sind die Bilder dagegen zu stark verrauscht, so ziehen sie auch die 'besseren' Bilder in Mitleidenschaft. Die Qualität der benutzten Bilder kann man mittels Epipolar- bzw. Trifokalgeometrie überprüfen.

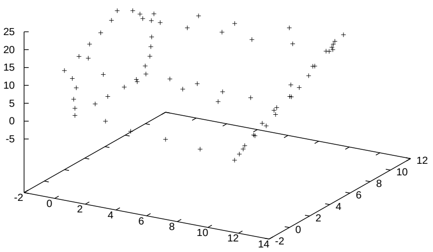
Die obigen Eigenschaften der SVD-Rekonstruktion sind unabhängig von der Weise, wie die projektiven Tiefen gesucht werden. Für die Experimente wurden allerdings die projektiven Tiefen benutzt, die mit Hilfe der Gleichung (4.4) gefunden wurden. Die Rekonstruktion anhand der Tiefen, die mit Invarianten gefunden wurden, ist aber nicht so stabil, da die Invarianten rauschempfindlich sind.



(*a*)



(*b*)



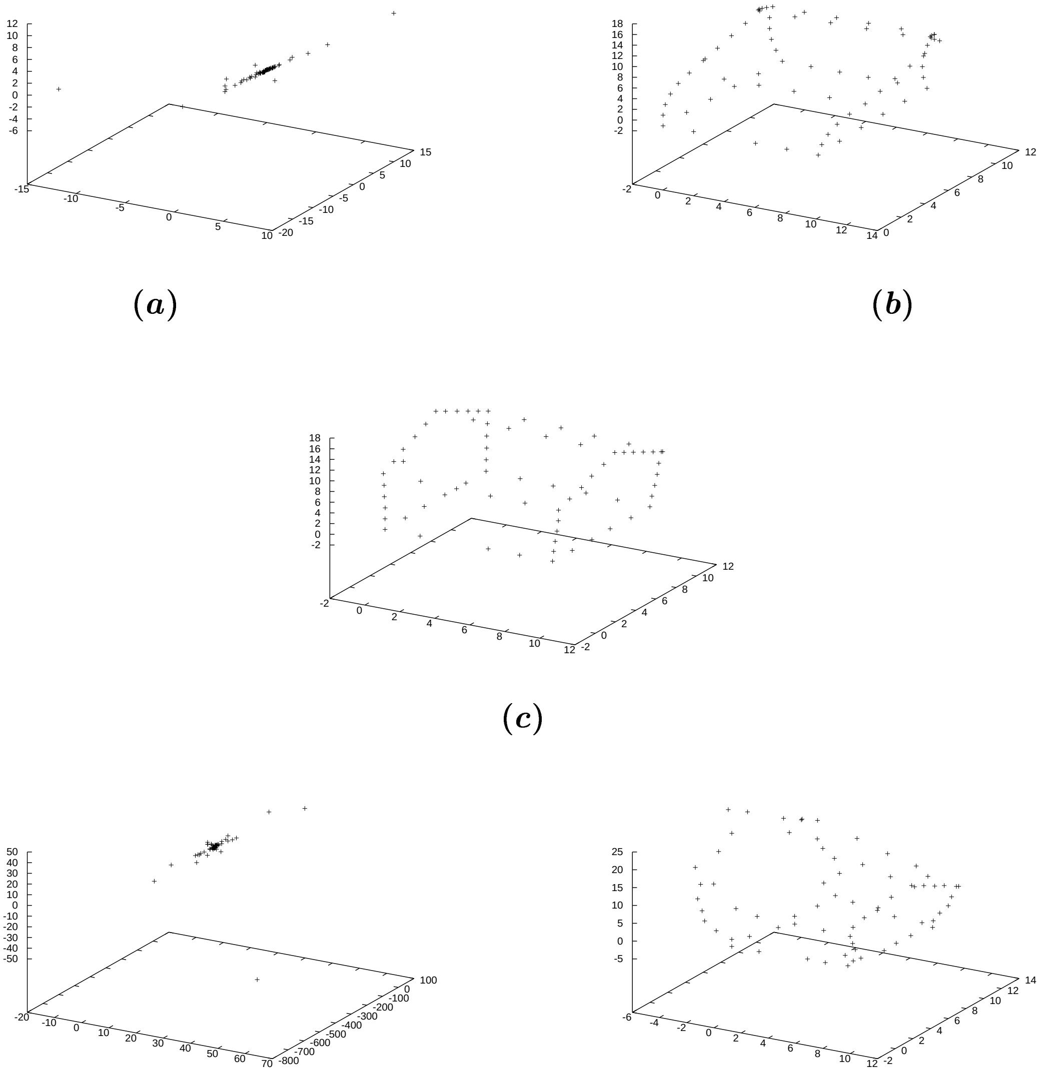
(*c*)

Abbildung 6.6: *Rekonstruktion mit den mittels Invarianten gefundenen Tiefen:* **(*a*)** *rauschfrei* **(*b*)** *0.01 Prozent Rauschen* **(*c*)** *0.05 Prozent Rauschen*

Die Abbildung 6.6 zeigt, daß im rauschfreien Fall die Ergebnisse gut sind, daß kleinstes Rauschen aber größere Einbußen in der Qualität des rekonstruierten Objektes verursacht. Im Verlauf des Experimentes wurde das Objekt (Haus) mit vier virtuellen Kameras aufgenommen und zunehmend verrauscht. Wie man der Abbildung 6.6 entnehmen kann, wirkt schon das Rauschen under 0.1 Prozent sehr stark auf die Rekonstruktion. Die Rekonstruktion ohne die Benutzung von Invarianten ist dagegen viel stabiler (siehe Abbildung 6.4).

6.3 Rekonstruktion mittels Invarianten

Im Gegensatz zur SVD-Rekonstruktion werden bei der Rekonstruktion mittels Invarianten die Punkte einzeln verarbeitet. Dieses bedeutet, daß nur die Qualität von einzelnen Korrespondenzen berücksichtigt werden muß. Sind einige Punkte sehr stark verrauscht, so werden nur diese 'schlecht' rekonstruiert, es wird aber kein Einfluß auf die anderen Punkte ausgeübt. Diese Art von Rekonstruktion eignet sich also besonders gut für die Objekte, bei denen die meisten Punkte rauschfrei, aber einige wenige sehr stark verrauscht sind. In der Abbildung 6.7 sind die Ergebnisse der Rekonstruktion mittels Invarianten repräsentiert.



(***a***) (***b***) (***c***) (***d***) (***e***)

Abbildung 6.7: *Rekonstruktion mittels Invarianten:* (***a***) *0.05 Prozent Rauschen mit den Fundamentalmatrizen* (***b***) *0.05 Prozent Rauschen mit dem Trifokaltensor* (***c***) *0.01 Prozent Rauschen bei SVD-Rekonstruktion* (***d***) *0.01 Prozent Rauschen mit den Fundamentalmatrizen* (***e***) *0.01 Prozent Rauschen mit dem Trifokaltensor*

In diesem Experiment wurde das Objekt in die vier virtuellen Kameras projiziert und mit 0.05-prozentigem Rauschen versehen. Danach wurde die Rekonstruktion mit Fundamentalmatrizen (6.7***a***) bzw. Trifokaltensoren (6.7***b***) durchgeführt. Das Resultat zeigt deutlich, daß das Rauschniveau von 0.05 Prozent schon große Schwierigkeiten für die Rekonstruktion bereitet. Deshalb wurde das Rauschen auf 0.01 Prozent heruntergesetzt. Außer der Invariantenrekonstruktion (6.7***d*** und 6.7***e***) wurde zum Vergleich auch SVD-Rekonstruktion mit denselben Daten durchgeführt (6.7***c***).

Es hat sich herausgestellt, daß die SVD-Rekonstruktion am stabilsten ist. Bei der Invariantenrekonstruktion ist die trifokale Geometrie der Stereogeometrie vorzuziehen, vorausgesetzt, daß die Korrespondenzen gut bestimmt sind.

6.4 Rekonstruktion mit echten Daten

Nachdem die Expermente mit künstlichen Bildern die Theorie bestätigt haben, wollen wir das Verhalten von verschiedenen Rekonstruktionsarten mit echten Daten näher untersuchen.

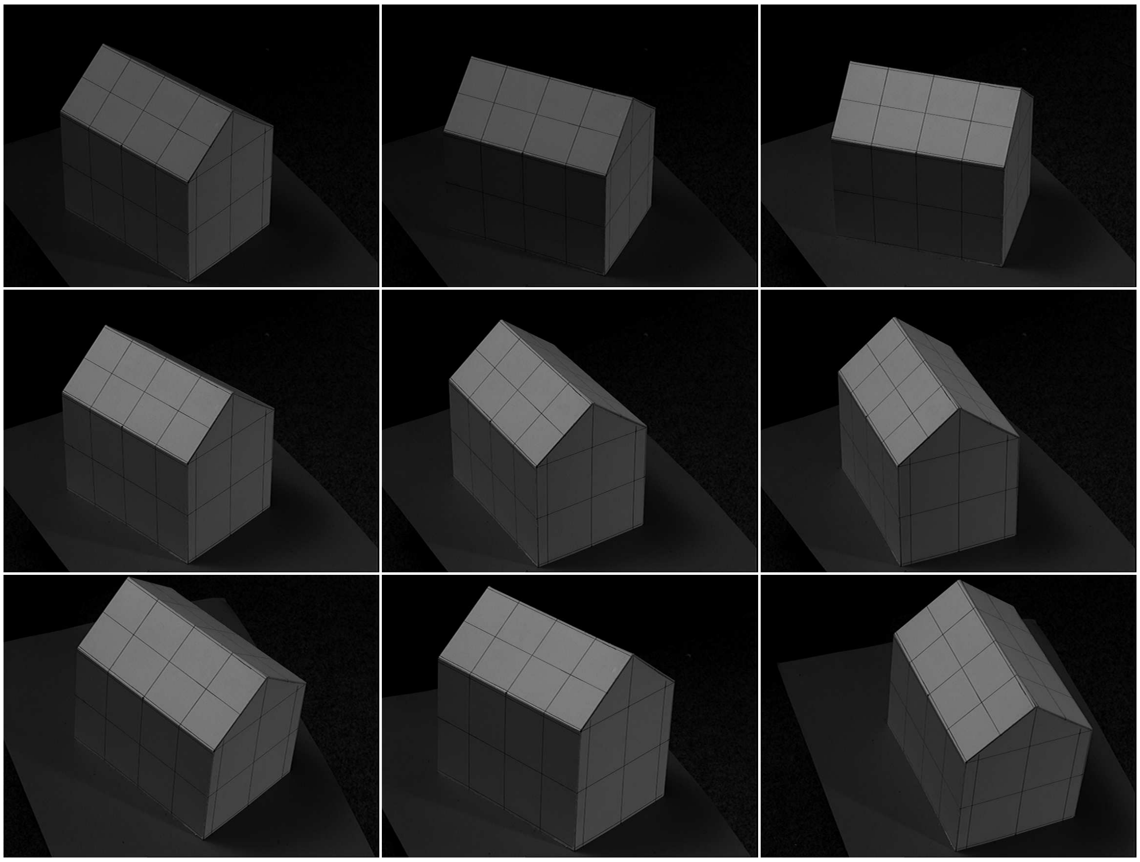


Abbildung 6.8: *Echte Daten*

Für die Experimente mit echten Daten wurden neun Bilder von einem Haus mit der Auflösung 512×512 Pixel aufgenommen (siehe Abbildung 6.8).

Um die Rekonstruktion gezielt beeinflussen und beurteilen zu können, wurden die

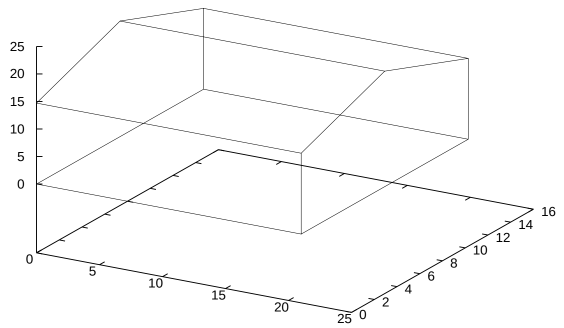


Abbildung 6.9: *Modell des Hauses*

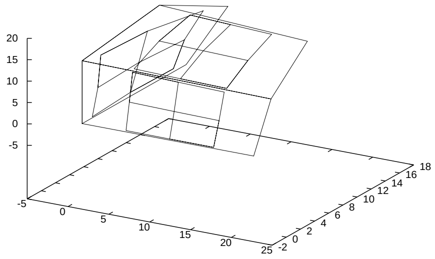
euklidischen Maße des Hauses im Raum ausgemessen (Abbildung 6.9).

Für die Korrespondenzen wurden die markanten Punkte (Eckpunkte und Schnittpunkte der Geraden) in jedem Bild manuell gewählt.

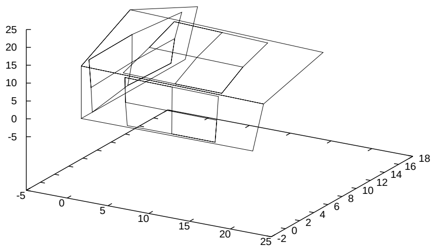
Die Epipolargeometrie wurde anhand dieser Korrespondenzen linear berechnet. Danach wurden Fundamentalmatrizen bezüglich der Frobeniusnorm minimiert, so daß die resultierenden Matrizen den Rang zwei hatten.

Die Invariantenrekonstruktion sowie die SVD-Rekonstruktion wurden mit diesen Daten getestet. Dabei hat sich herausgestellt, daß die auf den Invarianten basierenden Verfahren für die Rekonstruktion mit echten Daten (manuell ausgewählte Punkte) nicht geeignet sind. Dieses Ergebnis war aufgrund der Experimente mit künstlichen Bildern vorauszusehen – bei der manuellen Wahl der Punkte werden die Korrespondenzen ganzzahlig gesetzt, sodaß der durchschnittliche Fehler ein Pixel (bezüglich der Bilddimension $\frac{1}{512} \approx 0.2$ Prozent) beträgt. Bei solchem Rauschniveau versagt die Invariantenrekonstruktion. Dagegen funktioniert die SVD-Rekonstruktion bei den Bildern mit solchem Rauschen akzeptabel.

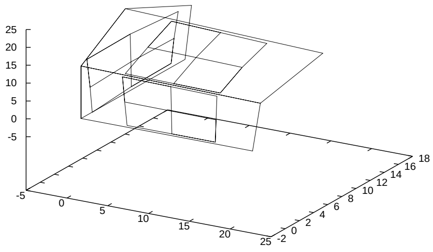
Abbildung 6.10 demonstriert die Ergebnisse dieser Rekonstruktion. Im Verlauf des Experimentes wurde eine zunehmende Anzahl von Bildern benutzt – zunächst wurden drei, dann sechs und schließlich neun Bilder im Experiment berücksichtigt. Den Erwartungen entsprechend hat sich die Qualität des rekonstruierten Objektes mit der Zunahme an Bildern verbessert. Wie man feststellen kann, spielt das Hinzufügen der Bilder besonders am Anfang eine sehr große Rolle – zwischen der Rekonstruktion mit drei und sechs Bildern gibt es größere Unterschiede als zwischen sechs und neun. Das ist darauf zurückzuführen, daß ab einer bestimmten Schwelle keine neuen Informationen durch die neuen Bilder gewonnen werden können.



(*a*)



(*b*)



(*c*)

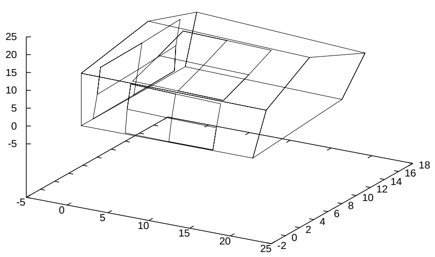
Abbildung 6.10: *Rekonstruiertes Haus:* **(*a*)** *mit 3 Bildern* **(*b*)** *mit 6 Bildern* **(*c*)** *mit 9 Bildern*

6.5 Erweiterte Rekonstruktion

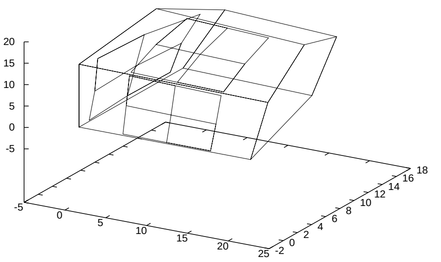
In diesem Abschnitt wollen wir die Möglichkeiten der erweiterten Rekonstruktion demonstrieren. Wir wollen beide Fälle der erweiterten Rekonstruktion behandeln – in der Bildebene und im Raum. Angesichts der Tatsache, daß die Rekonstruktion mittels Invarianten ziemlich instabil ist, wird nur die Rekonstruktion anhand des SVD-Algorithmuses mit echten Daten durchgeführt.

In der Abbildung 6.11 werden die Ergebnisse der erweiterten Rekonstruktion in der Bildebene sowie im Raum präsentiert. Das Experiment wurde mit den gleichen Parametern wie in Abschnitt 6.4 durchgeführt. Die Rekonstruktion wurde um zwei Eckpunkte erweitert, die in den aufgenommenen Bildern nicht zu sehen sind. Die Rekonstruktion wurde mit drei, sechs und neun Kameras in den Bildebenen sowie im Raum gemacht. Die linke und rechte Seite der Abbildung 6.11 zeigen dementsprechend Bild- und Raumrekonstruktion in Abhändigkeit von der Kameraanzahl.

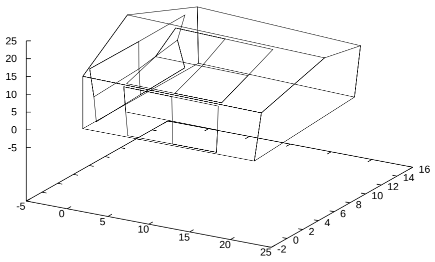
Die Ergebnisse zeigen deutlich, daß bei der kleinen Anzahl von Kameras die Bildrekonstruktion besser ist. Mit der Zunahme von Bildern gleichen sich aber die Ergebnisse an. Das ist darauf zurückzuführen, daß bei der geringeren Kameraanzahl das rekonstruierte Objekt nicht so 'gut' ist. Da man mit 'schlecht' rekonstruierten Punkten die Verdeckungen sucht, wirkt sich dieser Fehler auch auf die gefundenen verdeckten Punkte aus (6.11**a** bzw. 6.11**b**). Bei der großen Kameraanzahl spielt



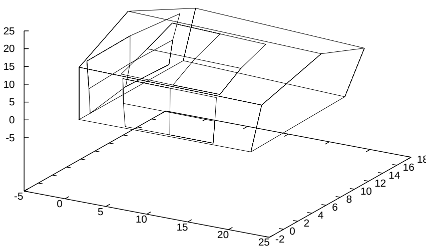
(a)



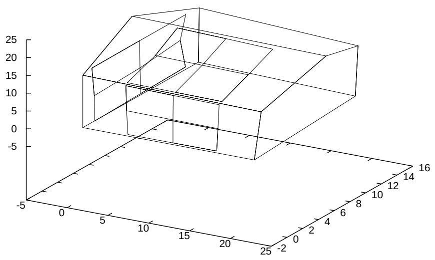
(b)



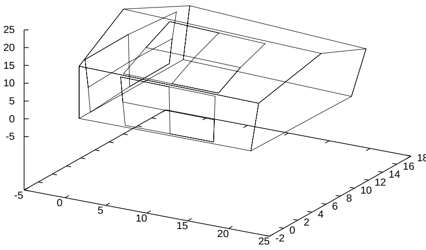
(c)



(d)



(e)



(f)

Abbildung 6.11: *Rekonstruierte verdeckte Punkte:* ***(a)*** *mit 3 Bildern in der Bildebene* ***(b)*** *mit 3 Bildern im Raum* ***(c)*** *mit 6 Bildern in der Bildebene* ***(d)*** *mit 6 Bildern im Raum* ***(e)*** *mit 9 Bildern in der Bildebene* ***(f)*** *mit 9 Bildern im Raum* dieser Fehler keine bedeutende Rolle mehr (6.11***e*** bzw. 6.11***f***).

Ein weiteres Merkmal basiert auf der batch-Eigenschaft der SVD-Rekonstruktion – das Einfügen von Punkten im Bild ändert die gesamte Rekonstruktion, das Einfügen von Punkten im Raum zeigt dagegen keine Auswirkung auf schon rekonstruierte Punkte (vergleiche Abbildungen 6.10 und 6.11).

6.6 Zusammenfassung

Die Experimente haben belegt, daß die SVD-Rekonstruktion der Rekonstruktion mittels Invarianten wegen des Verhaltens gegenüber dem Rauschen in den meisten Fällen vorzuziehen ist. Die Invariantenrekonstruktion ist dagegen nur bei geringem Rauschniveau einsetzbar. Zum Zweck der Bildanalyse und Rekonstruktion von verdeckten Punkten sind die Invarianten jedoch geeignet. Dabei ist die Koplanaritätsinvariante am ehesten anwendbar, da sie nur aus einem Bracket besteht. Alle anderen Invarianten setzen sich aus mehreren Brackets zusammen, wodurch deren Stabilität gegenüber dem Rauschen stark leidet.

Die Qualität des rekonstruierten Objektes hängt stark von der Qualität der Punktkorrespondenzen und der daraus berechneten Stereo- und Trifokalgeometrie ab.

Die projektiven Tiefen können zur Abschätzung des Abstandes zwischen dem Objekt und der Kamera eingesetzt werden.

Die Anzahl der benutzten Daten spielt eine wichtige Rolle bei jeder Art der Rekonstruktion – je mehr Punktkorrespondenzen und Kameras eingesetzt werden, desto bessere Ergebnisse werden erzielt.

Jede Art der hier vorgestellten Rekonstruktionsmethoden ist echtzeitfähig.

Kapitel 7

Schlußfolgerung

Am Ausgangspunkt dieser Arbeit stand die Frage nach unkalibrierter Rekonstruktion.

Zur Beschreibung der Lösungswege für dieses Problem haben wir die Geometrische Algebra gewählt. Hierfür haben wir eine Einführung in die Geometrische Algebra gegeben. Diese enthielt die Beschreibung der grundlegenden Konzepte, des Aufbaus von 2-D und 3-D projektiven Räumen, der geometrischen Verbindungsoperationen join und meet sowie der Modellierung von Objekten im Raum.

Das Problem der unkalibrierten Rekonstruktion erfordert Kenntnisse über die Kamerageometrie. Deshalb wurde das Modell der Lochkamera in Termen der Geometrischen Algebra vorgestellt. Weiterhin wurde die Geometrie von mehreren Kameras ausführlich erläutert. Dabei haben wir die Mittel zur Beschreibung der Stereo- und Trifokalgeometrie – Fundamentalmatrix und Trifokaltensor – in Notation der Geometrischen Algebra anschaulich geometrisch erklärt und umformuliert.

Ein weiterer Aspekt zur Behandlung des Rekonstruktionsproblems bilden die Invarianten. Mit deren Hilfe ist es möglich, die Rekonstruktion durchzuführen sowie die geometrischen Zusammenhänge der Szene zu erkennen und zu analysieren. Für diese Zwecke haben wir uns einen Einblick in die Invarianztheorie verschafft. Es wurden Wege zur Abschätzung von Punktkoordinaten in Termen einer gewählten projektiven Basis vorgestellt. Die Möglichkeiten der Erzeugung von Invarianten aus einer allgemeinen Invariantengleichung wurden beschrieben. Weiterhin haben wir gezeigt, wie man anhand der Bilddaten und der Kenntnisse der Epipolar- und Trifokalgeometrie diese mit lediglich gemessenen Daten berechnen kann.

Auf der beschriebenen Theorie aufbauend haben wir dann die verschiedenen bekannten Rekonstruktionsverfahren erklärt und erweitert. Zunächst wurde die SVD-Methode behandelt. Wir haben gesehen, welche Bedeutung die projektive Tiefe für diese Art der Rekonstruktion hat. Es wurden auch die verschiedenen Methoden zur Abschätzung von diesen Tiefen vorgestellt. Anschließend haben wir die Invarianztheorie für die Rekonstruktion eingesetzt. Weiterhin wurden die Möglichkeiten der Szeneanalyse mittels Geometrischer Algebra und Invarianten sowie der a priori Kenntnisse über die Umgebung für die Ergänzung der Szene beschrieben. Danach wurden die Wege zur Überführung einer projektiven in eine euklidische Rekonstruktion verdeutlicht.

Der Zweck der Rekonstruktion besteht darin, die Umgebung kennenzulernen, um mit dieser besser agieren zu können. Um bestimmte Manipulationen mit der Umgebung durchzuführen, bedarf man aber nicht immer einer Rekonstruktion. Dies haben wir anhand eines Beispiels demonstriert. Dabei handelte es sich um ein Greifproblem. Dieses Problem haben wir mit a priori Wissen über die Struktur der Szeneobjekte, nicht jedoch ihre geometrische Beziehungen zueinander unter dem Einsatz von Geometrischer Algebra und Invarianztheorie beispielhaft gelöst.

In Experimenten haben wir die hergeleiteten Rekonstruktonsmethoden ausgiebig getestet. Die Untersuchung der Robustheit dieser Verfahren spielte dabei eine große Rolle. Wir haben die Rekonstruktionsergebnisse unter variierender Anzahl von Punktkorrespondenzen und Kameras und variierendem Rauschniveau verglichen. Wie erwartet, waren die invariantenbasierten Methoden weniger robust als die Faktorisierungsmethoden. Die Rolle der projektiven Tiefe für die Abschätzung von der Entfernung zwischen den Objekten im Raum und den Kameras wurde mittels Experimenten erkärt. Es wurden auch die Eigenschaften der erweiterten Rekonstruktion untersucht.

Eine allgemeine Schwäche der punktbasierten Rekonstruktionen liegt in der relativ hohen Rauschempfindlichkeit. Diesem Problem liegt die Tatsache zugrunde, daß für solche Methoden die Kamerageometrie eingesetzt wird. Die Punkte werden eigentlich als die Schnittpunkte von optischen Strahlen berechnet. Daher wirken die kleinsten Abweichungen in der Bildebene sich sehr stark auf die Qualität der Rekonstruktion aus. Die regionbasierten Methoden sind dagegen viel stabiler. Allerdings sind solche Methoden meistens durch neuronalen Netze realisiert. Dies erfordert a priori Wissen über die zu rekonstruierenden Objekte. Oft sind neuronalen Netze auch nicht echtzeitfähig.

Dagegen sind alle hier vorgestellten Rekonstruktionsverfahren in Echtzeit ausführbar. Als Eingabedaten erwarten sie die Punktkorrespondenzen in mehreren Kameras. Genau dieser Schritt bereitet die meisten Probleme – das automatische Auffinden von Korrespondenzen benötigt den Einsatz von Eckendetektierungsmethoden, die oft sehr viel Zeit für die Berechnungen brauchen. Allerdings liefern solche Verfahren die Ergebnisse mit Pixel- oder sogar Subpixelgenauigkeit. Das erhöht entsprechend die Qualität der Rekonstruktion.

Um eine projektive in eine euklidische Rekonstruktion zu überführen, haben wir a priori Wissen über die Szene gebraucht. Dieses Wissen benötigten wir auch für einige Aspekte der erweiterten Rekonstruktion. Um die Analyse der Bilddaten zu beschleunigen und nicht einfach alle möglichen Kombinationen durchzutesten, bedarf es auch einiger Vorkenntnisse über die Umgebung. Die von uns vorgestellten Methoden sind nicht in der Lage diese Kentnisse selbst zu erwerben – sie erhalten diese als Eingabe. Um den ganzen Prozess zu automatisieren, sollte man an dieser Vorstufe nichtlineare analytische Werkzeuge einsetzen. Diese Aufgabe könnte z.B.

durch neuronalen Netze bearbeitet werden.

Wir halten als wichtige Ergebnisse fest:

- Die Nutzung der Geometrischen Algebra ermöglicht eine einfache Beschreibung der Szenengeometrie.
- Die Invarianten beschreiben übersichtlich die geometrischen Zusammenhänge zwischen Objekten im Raum und in der Bildebene.
- Die projektiven Tiefen lassen sich einfach mittels Invarianten berechnen, wenn die entsprechende Raumbasis gewählt wurde.
- Die Kenntnisse über die projektive Tiefe kann man für die Abschätzung der Entfernung des Objektes zum Beobachter einsetzen.
- Invarianztheorie bietet die Möglichkeit der projektiven Rekonstruktion von Gestalt sowie von der Bewegung.
- Die Faktorisierungsmethoden sind bei der Rekonstruktion robuster als Invarianenmethoden.
- Durch die Bildanalyse mittels Geometrischer Algebra und der Invarianztheorie läßt sich die Szene um verdeckte und neue Punkte erweitern.
- Ohne des a priori Wissens ist es nicht möglich, die Szene effektiv zu ergänzen und euklidisch zu rekonstruieren.

Anhang A

Kameramodell

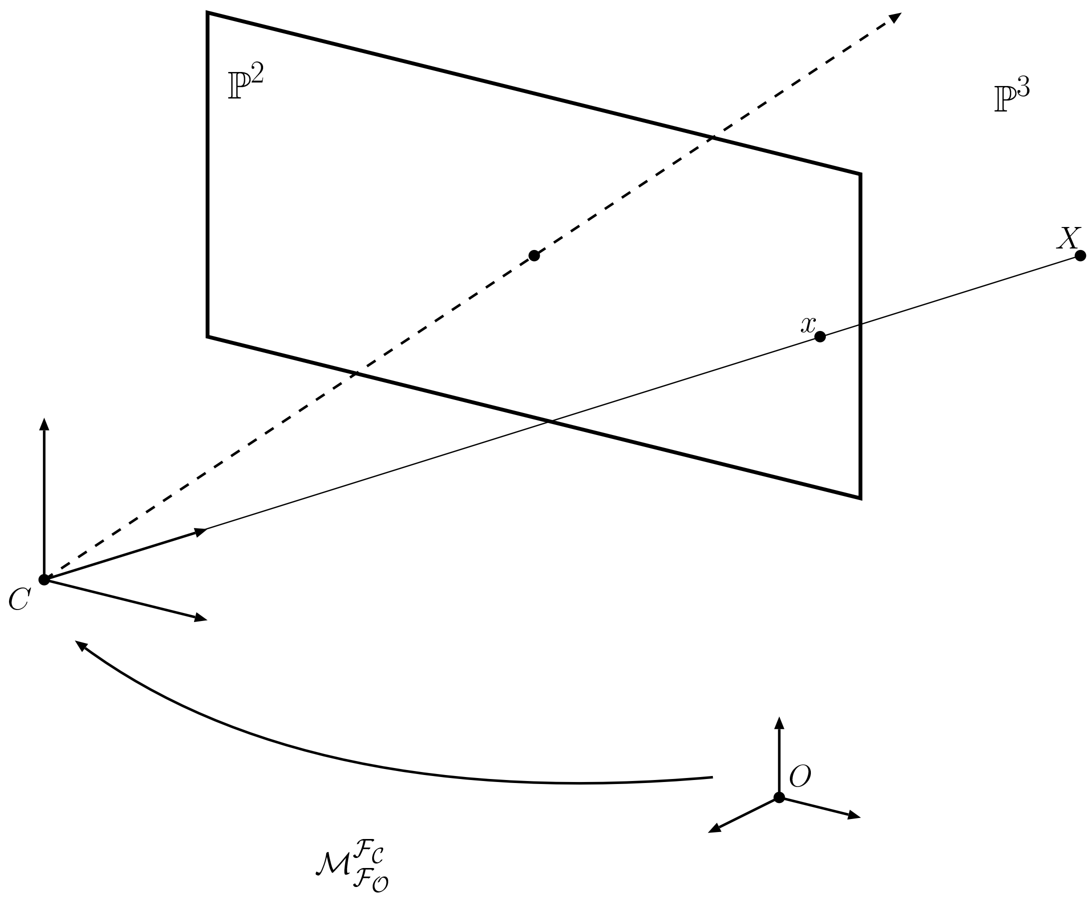


Abbildung A.1: *Lochkameramodell*

In Abbildung A.1 wird die Projektion eines Punktes $\boldsymbol{X}$ aus $\mathbb{P}^3$ in die Bildebene dargestellt. Der abgebildete Punkt $\boldsymbol{x}$ liegt in $\mathbb{P}^2$. Das optische Zentrum der Kamera wird mit $\boldsymbol{C}$ bezeichnet. Der Koordinatenursprung von $\mathbb{P}^3$ liegt in $\boldsymbol{O}$.

Die Projektion ist eine lineare Abbildung von $\mathbb{P}^3$ in $\mathbb{P}^2$ und kann mit Hilfe einer $\mathbf{3} \times \mathbf{4}$ Matrix $\boldsymbol{P}$ beschrieben werden:

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{X}.$$

In Abhängigkeit von der Distanz zwischen dem betrachteten Objekt und dem optischen Zentrum der Kamera, kann man drei Kameramodelle unterscheiden.

Ist das Objekt der Kamera nah, so wird das *Lochkamera*-Modell genommen. Dieses Modell ist durch die Projektionsmatrix

$$\boldsymbol{P} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{14} \\ t_{31} & t_{32} & t_{33} & t_{34} \end{bmatrix} \underbrace{\begin{bmatrix} \sigma f & -\cot\alpha & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{K_p}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\boldsymbol{P_0}} \boldsymbol{M}_{\boldsymbol{F_0}}^{\boldsymbol{F_C}} \tag{A.1}$$

definiert, wobei $\boldsymbol{K_p}$ die Transformation in der Bildebene, die sogenannten *intrinsischen Parameter*, beschreibt:

- Die *Brennweite* $\boldsymbol{f}$ – der Abstand zwischen dem optischen Zentrum und der Bildebene
- Das *Bildseitenverhältnis* (*Aspectratio*) $\boldsymbol{\sigma}$ – Skalierungsfaktor zwischen den Koordinatenrichtungen in der Bildebene
- $\boldsymbol{\alpha}$ – der Winkel zwischen den Koordinatenrichtungen
- $(\boldsymbol{u_0}, \boldsymbol{v_0})$ – der Schnittpunkt der optischen Achse mit der Bildebene

Die $\boldsymbol{M}_{\boldsymbol{F_0}}^{\boldsymbol{F_C}}$-Matrix beschreibt die Transformation zwischen den Weltkoordinaten $\boldsymbol{\mathcal{F}_0}$ und Kamerakoordinaten $\boldsymbol{\mathcal{F}_C}$:

$$\boldsymbol{M}_{\boldsymbol{F_0}}^{\boldsymbol{F_C}} = \begin{bmatrix} \boldsymbol{R} & \vec{\boldsymbol{t}} \\ \vec{\boldsymbol{0}}^{\boldsymbol{T}} & 1 \end{bmatrix}$$

mit der $\boldsymbol{3} \times \boldsymbol{3}$ Rotationsmatrix $\boldsymbol{R}$, dem $\boldsymbol{3} \times \boldsymbol{1}$ Translationsvektor $\vec{\boldsymbol{t}}$ und dem $\boldsymbol{3} \times \boldsymbol{1}$ Nullvektor $\vec{\boldsymbol{0}}$.

Ist das Objekt weit von der Kamera entfernt, so sind die optischen Strahlen (fast) parallel. In diesem Fall spricht man von einer *affinen Kamera*. Diese kann man sich als eine Lochkamera vorstellen, deren optisches Zentrum im Unendlichen liegt:

$$\boldsymbol{P} = \underbrace{\begin{bmatrix} \sigma f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{K_{aff}}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{P_{par}}} \boldsymbol{M}_{\boldsymbol{F_0}}^{\boldsymbol{F_C}}, \tag{A.2}$$

wobei $\boldsymbol{P_{par}}$ die parallele Projektion beschreibt.

Spielen die Skalierungsfaktoren keine Rolle, so wird aus der affinen die *orthografische Kamera*:

$$\boldsymbol{P} = \underbrace{\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}}_{\boldsymbol{K_{ort}}} \underbrace{\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}}_{\boldsymbol{P_{par}}} \boldsymbol{M}_{\boldsymbol{F_0}}^{\boldsymbol{F_C}}. \tag{A.3}$$

Anhang B

Gleichung für vier Ebenen

Sei $\boldsymbol{\Phi}^{\boldsymbol{A}}_{\boldsymbol{ij}}$ die Ebene, die durch die Punkte $\boldsymbol{A}'_1$,$\boldsymbol{A}'_2$ in der Bildebene und das optische Zentum $\boldsymbol{A_0}$ der Kamera $\boldsymbol{A}$ geht. Entsprechend sei $\boldsymbol{\Phi}^{\boldsymbol{B}}_{\boldsymbol{ij}}$ für Kamera $\boldsymbol{B}$ definiert. Dann läßt sich folgende Gleichung aufstellen:

$$\begin{aligned}
\boldsymbol{X}_1 \bigwedge \boldsymbol{X}_2 \bigwedge \boldsymbol{X}_3 \bigwedge \boldsymbol{X}_4 &= (\boldsymbol{X}_1 \bigwedge \boldsymbol{X}_2) \bigwedge (\boldsymbol{X}_3 \bigwedge \boldsymbol{X}_4) \\
&= ((\boldsymbol{A}_0 \bigwedge \boldsymbol{A}'_1 \bigwedge \boldsymbol{A}'_2) \bigvee (\boldsymbol{B}_0 \bigwedge \boldsymbol{B}'_1 \bigwedge \boldsymbol{B}'_2)) \bigwedge \qquad \text{(B.1)} \\
&\quad ((\boldsymbol{A}_0 \bigwedge \boldsymbol{A}'_3 \bigwedge \boldsymbol{A}'_4) \bigvee (\boldsymbol{B}_0 \bigwedge \boldsymbol{B}'_3 \bigwedge \boldsymbol{B}'_4)) \\
&= (\boldsymbol{\Phi}^{\boldsymbol{A}}_{\boldsymbol{12}} \bigvee \boldsymbol{\Phi}^{\boldsymbol{B}}_{\boldsymbol{12}}) \bigwedge (\boldsymbol{\Phi}^{\boldsymbol{A}}_{\boldsymbol{34}} \bigvee \boldsymbol{\Phi}^{\boldsymbol{B}}_{\boldsymbol{34}}).
\end{aligned}$$

Mit (1.20) und der Tatsache, daß $\boldsymbol{\mathcal{I}^2 = 1}$ ist, ergibt sich:

$$\begin{aligned}
[\mathbf{1234}] &= ((\boldsymbol{\Phi}^{\boldsymbol{A}}_{\boldsymbol{12}} \bigvee \boldsymbol{\Phi}^{\boldsymbol{B}}_{\boldsymbol{12}}) \bigwedge (\boldsymbol{\Phi}^{\boldsymbol{A}}_{\boldsymbol{34}} \bigvee \boldsymbol{\Phi}^{\boldsymbol{B}}_{\boldsymbol{34}}))\boldsymbol{\mathcal{I}}^{-1} \\
&= (\{(\boldsymbol{\Phi}^{\boldsymbol{A}\,*}_{\boldsymbol{12}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{B}\,*}_{\boldsymbol{12}}\}\boldsymbol{\mathcal{I}}) \bigwedge (\{\boldsymbol{\Phi}^{\boldsymbol{A}\,*}_{\boldsymbol{34}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{B}\,*}_{\boldsymbol{34}}\}\boldsymbol{\mathcal{I}}))\boldsymbol{\mathcal{I}}^{-1} \\
&= (\boldsymbol{\Phi}^{\boldsymbol{A}\,*}_{\boldsymbol{12}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{B}\,*}_{\boldsymbol{12}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{A}\,*}_{\boldsymbol{34}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{B}\,*}_{\boldsymbol{34}})\boldsymbol{\mathcal{I}}^{-1} \\
&= -(\boldsymbol{\Phi}^{\boldsymbol{A}\,*}_{\boldsymbol{12}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{A}\,*}_{\boldsymbol{34}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{B}\,*}_{\boldsymbol{12}} \bigwedge \boldsymbol{\Phi}^{\boldsymbol{B}\,*}_{\boldsymbol{34}})\boldsymbol{\mathcal{I}}^{-1} \\
&= -((\boldsymbol{\Phi}^{\boldsymbol{A}}_{\boldsymbol{12}} \bigvee \boldsymbol{\Phi}^{\boldsymbol{A}}_{\boldsymbol{34}}) \bigwedge (\boldsymbol{\Phi}^{\boldsymbol{B}}_{\boldsymbol{12}} \bigvee \boldsymbol{\Phi}^{\boldsymbol{B}}_{\boldsymbol{34}}))\boldsymbol{\mathcal{I}}^{-1}. \qquad \text{(B.2)}
\end{aligned}$$

Anhang C

Korrekturkoeffizienten für Invarianten

Der Schnittpunkt $\boldsymbol{a}'_{ijkl}$ von zwei Geraden $\boldsymbol{a}'_i \bigwedge \boldsymbol{a}'_j$ und $\boldsymbol{a}'_k \bigwedge \boldsymbol{a}'_l$ kann als lineare Kombination von $\boldsymbol{a}'_k$ und $\boldsymbol{a}'_l$ dargestellt werden:

$$\boldsymbol{a}'_{ijkl} = \boldsymbol{\mu}_{ijkl}\boldsymbol{a}'_l + (\mathbf{1} - \boldsymbol{\mu}_{ijkl})\boldsymbol{a}'_k. \tag{C.1}$$

Sind $\boldsymbol{A}'_i$,$\boldsymbol{A}'_j$,$\boldsymbol{A}'_k$,$\boldsymbol{A}'_l$ die $\mathbb{P}^3$-Repräsentationen von den Punkten $\boldsymbol{a}'_i$,$\boldsymbol{a}'_j$,$\boldsymbol{a}'_k$,$\boldsymbol{a}'_l$, so gilt:

$$\begin{aligned}
\boldsymbol{A}'_{ijkl} &= (\boldsymbol{A}'_i \bigwedge \boldsymbol{A}'_j) \bigvee (\boldsymbol{A}_0 \bigwedge \boldsymbol{A}'_k \bigwedge \boldsymbol{A}'_l) \\
&= [\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}_0\boldsymbol{A}'_k]\boldsymbol{A}'_l + [\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_l\boldsymbol{A}_0]\boldsymbol{A}'_k + [\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_k\boldsymbol{A}'_l]\boldsymbol{A}_0.
\end{aligned}$$

Da $\boldsymbol{A}'_i$,$\boldsymbol{A}'_j$,$\boldsymbol{A}'_k$,$\boldsymbol{A}'_l$ koplanar sind, ist $[\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_k\boldsymbol{A}'_l] = \mathbf{0}$, und somit

$$\boldsymbol{A}'_{ijkl} = [\boldsymbol{A}_0\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_k]\boldsymbol{A}'_l - [\boldsymbol{A}_0\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_l]\boldsymbol{A}'_k. \tag{C.2}$$

Aus (C.2) ergibt sich:

$$\begin{aligned}
\boldsymbol{a}'_{ijkl} &= \frac{\boldsymbol{A}'_{ijkl} \wedge \boldsymbol{\gamma}_4}{\boldsymbol{A}'_{ijkl} \cdot \boldsymbol{\gamma}_4} \\
&= [\boldsymbol{A}_0\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_k]\frac{\boldsymbol{A}'_l \wedge \boldsymbol{\gamma}_4}{\boldsymbol{A}'_{ijkl} \cdot \boldsymbol{\gamma}_4} - [\boldsymbol{A}_0\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_l]\frac{\boldsymbol{A}'_k \wedge \boldsymbol{\gamma}_4}{\boldsymbol{A}'_{ijkl} \cdot \boldsymbol{\gamma}_4} \\
&= [\boldsymbol{A}_0\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_k]\frac{\boldsymbol{A}'_l \cdot \boldsymbol{\gamma}_4}{\boldsymbol{A}'_{ijkl} \cdot \boldsymbol{\gamma}_4}\boldsymbol{a}'_l - [\boldsymbol{A}_0\boldsymbol{A}'_i\boldsymbol{A}'_j\boldsymbol{A}'_l]\frac{\boldsymbol{A}'_k \cdot \boldsymbol{\gamma}_4}{\boldsymbol{A}'_{ijkl} \cdot \boldsymbol{\gamma}_4}\boldsymbol{a}'_k,
\end{aligned} \tag{C.3}$$

wobei

$$\boldsymbol{a}'_l = \frac{\boldsymbol{A}'_l \wedge \boldsymbol{\gamma}_4}{\boldsymbol{A}'_{ijkl} \cdot \boldsymbol{\gamma}_4} \qquad \boldsymbol{a}'_k = \frac{\boldsymbol{A}'_k \wedge \boldsymbol{\gamma}_4}{\boldsymbol{A}'_{ijkl} \cdot \boldsymbol{\gamma}_4}.$$

Vergleicht man das obige Ergebnis mit (C.1), so gilt:

$$\begin{aligned}
\boldsymbol{\mu_{ijkl}} &= \frac{\boldsymbol{A'_l} \cdot \gamma_4}{\boldsymbol{A'_{ijkl}} \cdot \gamma_4}[\boldsymbol{A_0 A'_i A'_j A'_k}] \\
\boldsymbol{\mu_{ijkl}} - \mathbf{1} &= \frac{\boldsymbol{A'_k} \cdot \gamma_4}{\boldsymbol{A'_{ijkl}} \cdot \gamma_4}[\boldsymbol{A_0 A'_i A'_j A'_l}].
\end{aligned}$$

Auf diese Weise kann man die Korrekturkoeffizienten für Invarianten berechnen. So gilt z.B.:

$$\begin{aligned}
\frac{\boldsymbol{\mu_{1245}\mu_{1346}}}{\boldsymbol{\mu_{1246}\mu_{1345}}} &= \frac{\frac{\boldsymbol{A'_5}\cdot\gamma_4}{\boldsymbol{A'_{1245}}\cdot\gamma_4}[\boldsymbol{A_0A'_1A'_2A'_4}]\frac{\boldsymbol{A'_6}\cdot\gamma_4}{\boldsymbol{A'_{1346}}\cdot\gamma_4}[\boldsymbol{A_0A'_1A'_3A'_4}]}{\frac{\boldsymbol{A'_6}\cdot\gamma_4}{\boldsymbol{A'_{1246}}\cdot\gamma_4}[\boldsymbol{A_0A'_1A'_2A'_4}]\frac{\boldsymbol{A'_5}\cdot\gamma_4}{\boldsymbol{A'_{1346}}\cdot\gamma_4}[\boldsymbol{A_0A'_1A'_3A'_4}]} \\
&= \frac{(\boldsymbol{A'_{1245}}\cdot\gamma_4)(\boldsymbol{A'_{1346}}\cdot\gamma_4)}{(\boldsymbol{A'_{1246}}\cdot\gamma_4)(\boldsymbol{A'_{1345}}\cdot\gamma_4)}.
\end{aligned}$$

Für Kamera $\boldsymbol{B}$ kann man $\boldsymbol{\nu_{ijkl}}$ definieren:

$$\boldsymbol{b'_{ijkl}} = \boldsymbol{\nu_{ijkl} b'_l} + (\mathbf{1} - \boldsymbol{\nu_{ijkl}})\boldsymbol{b'_k}. \tag{C.4}$$

Analog zu Kamera $\boldsymbol{A}$ gilt:

$$\frac{\boldsymbol{\nu_{1245}\nu_{1346}}}{\boldsymbol{\nu_{1246}\nu_{1345}}} = \frac{(\boldsymbol{B'_{1245}}\cdot\gamma_4)(\boldsymbol{B'_{1346}}\cdot\gamma_4)}{(\boldsymbol{B'_{1246}}\cdot\gamma_4)(\boldsymbol{B'_{1345}}\cdot\gamma_4)}.$$

Insgesamt ergibt sich:

$$\frac{\boldsymbol{\phi_{1245}\phi_{1346}}}{\boldsymbol{\phi_{1246}\phi_{1345}}} = \frac{\boldsymbol{\mu_{1245}\mu_{1346}\nu_{1245}\nu_{1346}}}{\boldsymbol{\mu_{1246}\mu_{1345}\nu_{1246}\nu_{1345}}}.$$

Anhang D

Acht-Punkte-Algorithmus

In [16] hat R.Hartley beschrieben, wie man mit acht Punktkorrespondenzen die Fundamentalmatrix finden kann. Dem Algorithmus zugrunde liegt folgende Gleichung:

$$\boldsymbol{x'^T F x = 0}.$$

Schreibt man die obige Gleichung wie folgt

$$(\mathrm{x'y'w'})\begin{pmatrix} \boldsymbol{f_{11}} & \boldsymbol{f_{12}} & \boldsymbol{f_{13}} \\ \boldsymbol{f_{21}} & \boldsymbol{f_{22}} & \boldsymbol{f_{23}} \\ \boldsymbol{f_{31}} & \boldsymbol{f_{32}} & \boldsymbol{f_{33}} \end{pmatrix}\begin{pmatrix} \mathrm{x} \\ \mathrm{y} \\ \mathrm{w} \end{pmatrix} = 0,$$

so gilt:

$$\begin{aligned} \mathrm{xx'}\boldsymbol{f_{11}} + \mathrm{yx'}\boldsymbol{f_{12}} + \mathrm{wx'}\boldsymbol{f_{13}} &\quad + \\ \mathrm{xy'}\boldsymbol{f_{21}} + \mathrm{yy'}\boldsymbol{f_{22}} + \mathrm{wy'}\boldsymbol{f_{23}} &\quad + \\ \mathrm{xw'}\boldsymbol{f_{31}} + \mathrm{yw'}\boldsymbol{f_{32}} + \mathrm{ww'}\boldsymbol{f_{33}} &= 0. \end{aligned} \tag{D.1}$$

Also liefert jede Punktkorrespondenz eine Gleichung, die in Koeffizienten der Fundamentalmatrix linear ist. Da die Fundamentalmatrix bis auf einen Skalarfaktor definiert ist, reichen acht Punktkorrespondenzen, um sie zu bestimmen.

Weil die echten Daten verrauscht sind, ist die auf die oben beschriebene Weise gefundene Fundamentalmatrix nicht exakt. Normalerweise hat die $\boldsymbol{F}$-Matrix den Rang 2. So wird die mit den echten Daten gefundene (verrauschte) Fundamentalmatrix zu der bezüglich der Frobeniusnorm nächsten Rang-Zwei-Matrix gebracht. Das kann man z.B. mit Hilfe von SVD machen, indem man den dritten Singulärwert der Fundamentalmatrix gleich Null setzt.

Analog zur Fundamentalmatrix kann man auch den Trifokaltensor bestimmen. Die dabei zugrunde liegenden Gleichungen sind:

$$\boldsymbol{T_{ABC} x_A l_B l_C = 0},$$

oder

$$\boldsymbol{l_A} = \boldsymbol{T_{ABC}} \boldsymbol{l_B} \boldsymbol{l_C}.$$

Anhang E

Wahl der projektiven Basis

Man kann fünf beliebige Punkte im $\mathbb{P}^3$ als eine projektive Basis des Raumes wählen, wenn jede Viererkombination dieser Punkte in allgemeiner Position (d.h. linear unabhängig) sind. Da man diese Wahl oft in der Bildebene macht, soll ein Test gefunden werden, der anhand der Bildpunkte durchgeführt werden kann. Eine solche Möglichkeit bietet das Bracket. Lineare Unabhängigkeit von vier Punkten ist mit der Aussage äquivalent, daß diese Punkte nicht koplanar sind.

Seien $\boldsymbol{\{X_i\}}$,$\boldsymbol{i} \in \boldsymbol{\{1, ..., 5\}}$ fünf Punkte im Raum und $\boldsymbol{\{x_{A_i}, x_{B_i}, x_{C_i}\}}$ die Abbildungen dieser Punke in den Kameras $\boldsymbol{A}$,$\boldsymbol{B}$ und $\boldsymbol{C}$. Sollen $\boldsymbol{X_i}$,$\boldsymbol{X_j}$,$\boldsymbol{X_k}$,$\boldsymbol{X_l}$ linear unabhängig sein, so gilt:

$$[\boldsymbol{X_i X_j X_k X_l}] \neq 0. \tag{E.1}$$

Mit Hilfe von (3.13) und (3.23) kann man (E.1) in Termen der Bildpunkte schreiben:

$$\boldsymbol{x^T_{B_{ijkl}} F_{AB} x_{A_{ijkl}}} \neq \boldsymbol{0}$$

und

$$\boldsymbol{T_{ABC} x_{A_{ijkl}} l_{B_{ij}} l_{C_{kl}}} \neq \boldsymbol{0},$$

wobei die Korrekturkoeffizenten keine Rolle spielen, da mit Null verglichen wird.

Gilt die Bedingung (E.1) für alle Kombinationen $\boldsymbol{(i, j, k, l)} \in \boldsymbol{\{1, ..., 5\}}$, so können die gewählten Punkte eine projektive Basis des Raumes bilden.

Literaturverzeichnis

[1] Shashua, A. Projective structure from uncalibrated images: structure from motion and recognition. IEEE Transactions on *Pattern Analysis and Machine Intelligence*, 16(8), pp. 778–790, 1994.

[2] Sparr, G. Kinetic depth. In Proceedings of *European Conference of Computer Vision*, Vol.2, pp. 471–482, 1994.

[3] Csurka, G. and Faugeras, O. Computing three dimensional project invariants from a pair of images using the Grassmann-Caley algebra. *Journal of Image and Vision Computing*, 16, pp. 3–12, 1998.

[4] Tomasi, C. and Kanade, T. Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, Vol.9(2), pp. 137–154, 1992.

[5] Poelman, C.J. and Kanade, T. A paraperspective factorization method for shape and motion recovery. In Proceedings of *European Conference of Computer Vision*, Stockholm, pp. 97–108, 1994.

[6] Triggs, W. Matching Constraints and the Joint Image. In Proceedings of *ICCV'95*, IEEE Computer Society Press, Boston, pp. 338–343, 1995.

[7] Hestenes, D. and Sobczyk, G. Clifford Algebra to Geometric Calculus: A unified language for mathematics and physics. *D. Reidel*, Dordrecht, 1984

[8] Hestenes, D. and Ziegler, R. Projective geometry with Clifford algebra. *Acta Applicandae Mathematicae*, Vol.23, pp. 25–63, 1991

[9] Grassmann, H. Der Ort der Hamilton'schen Quaternionen in der Ausdehnungslehre. *Math. Ann.*, 12: pp. 375, 1877

[10] Clifford, W.K. Applications of Grassmann's extensive algebra. *Am. J. Math.* 1: pp. 350–358, 1878.

[11] Hestenes, D. The design of linear algebra and geometry. *Acta Applicandae Mathematicae*, 23: pp. 65–93, 1991.

[12] Banarer, V. Grundlagen der Analyse des visuellen 3D-Raumes im Rahmen der Geometrischen Algebra. *Studienarbeit*, Institut für Informatik der Christian-Albrechts-Universität Kiel, 1998.

[13] Bayro-Corrochano E., Lasenby J., Sommer G. 1996. Geometric Algebra: A framework for computing point and line correspondences and projective structure using n uncalibrated cameras. In Proceedings of *ICPR*, Viena, Austria, Vol.1, pp. 334–338, 1996

[14] Bayro-Corrochano, E. and Lasenby, J. Geometric techniques for the computation of projective invariants using n uncalibrated cameras. In Proceedings of *Indian Conference on Computer Vision and Image Processing*, New Delphi, India, pp. 95–100, 1998

[15] Sturm, P. and Triggs, W. A factorization based algorithm for multi-image, projective structure and motion. In Proceedings of *European Conference of Computer Vision*, LNCS 1064, Springer Verlag, Berlin, Vol.2, pp. 709–720, 1996.

[16] Hartley, R. In defence of the 8-point algorithm. IEEE Transactions on *Pattern Analysis and Machine Intelligence*, Vol.19(6), pp. 580–593, 1997.

[17] Heyden, A. Geometry and Algebra of Multiple Projective Transformations. *Doctoral Dissertation*, 1995.

[18] Boufama, B. and Mohr, R. Epipole and fundamantal matrix estimation using virtual parallax. In Proceedings of *ICCV'95*, IEEE Computer Society Press, Boston, pp. 1030–1036, 1995.

[19] Carlsson, S. Symmetry in perspective. In Proceedings of *European Conference on Computer Vision*, Freiburg, Germany, pp. 249–263, 1998.

Erklärung

Hiermit versichere ich, daß ich die vorliegende Arbeit selbstständig verfaßt und ausschließlich die angegebenen Hilfsmittel und Quellen verwendet habe.

Kiel, den 20.05.1999