DISSERTATION

Geometry, Constraints and Computation of the Trifocal Tensor

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften unter Leitung von

Prof. Dr.-Ing. Wolfgang Förstner
Institut für Photogrammetrie, Universität Bonn

in Zusammenarbeit mit

O.Univ.Prof. Dipl.-Ing. Dr.techn. Karl Kraus
Institut für Photogrammetrie und Fernerkundung (E122),
Technische Universität Wien

eingereicht an der Technische Universität Wien
Fakultät für Naturwissenschaften und Informatik

von

Dipl.-Ing. Camillo Ressl
Matr.Nr 8926567
Flurschützstrasse 1/21
A-1120 Wien

Wien, im Juni 2003 .....

Diese Arbeit ist Teil des Forschungsprojekts „Bildorientierung in verschiedenen Disziplinen“ und wurde vom Fonds zur Förderung der wissenschaftlichen Forschung (FWF) unter Projekt-Nr. P13901-INF unterstützt.

This work is part of the research project “Image orientation in different disciplines” and was supported by the Austrian Science Fund under project no. P13901-INF.

Abstract

The topic of this thesis is the **trifocal tensor**, which describes the relative orientation (or epipolar geometry) of *three uncalibrated images*. So it plays the same role for three images as the fundamental matrix plays for two. The trifocal tensor is a homogenous valence-(1,2) tensor, which means that it can be represented as a $3\times3\times3$ cube of numbers. It is of particular interest because of the following properties:

– The trifocal tensor can be determined *linearly* from corresponding points *and* lines in three images; the latter are not useable for the fundamental matrix. Consequently the trifocal tensor provides *a tool to determine the relative orientation of three images without requiring approximate values*. This orientation can be used to initialize a subsequent bundle block adjustment. Furthermore, the chances of running into critical configurations for determining the trifocal tensor is much smaller than for two images and the fundamental matrix. The only practically relevant critical configuration for computing the trifocal tensor happens if all corresponding image features arise from a common plane in space.

  After determining the trifocal tensor, the basis vectors and rotation matrices of the relative orientation of the three images can be extracted easily if the interior orientation of the images is known. If the interior orientation is unknown, but the same for all three images, then, in general, this common interior orientation can be retrieved also.

  If $n > 3$ images are given, then many unrelated trifocal tensors are computed. Before all n images can be used simultaneously, the individual basis vectors and rotation matrices derived from the unrelated tensors must be transformed into one common system.

– The *transfer relations* associated with the trifocal tensor can be used to transfer the content (points and lines) of two source images into the other one (the target image). If the target image is a *real* image, the transferred position of features can be used to initialize a matching technique in the target image. If the target image is *virtual* with a given orientation, then the transfer relations can be used to form the content of this new image by transferring pixel by pixel. This process is called *novel view synthesis* or *image formation*.

In this thesis we concentrate on the first issue of determining the trifocal tensor.

Each triple of points provides 4 equations, called *trilinearities*, which are linear in the tensor elements and each corresponding triple of lines provides 2 linear equations. Since the tensor is made up of 27 elements, at least 7 points, 13 lines or a proper combination are needed for a direct linear solution of the trifocal tensor.

There are, however, some drawbacks associated with this linear solution. The trifocal tensor is made up of 27 elements, but it only has 18 degrees of freedom. Consequently its elements have to satisfy 8 internal constraints, besides the fixing of the scale ambiguity, to represent a *valid* trifocal tensor. These constraints are in generally not satisfied by the direct linear solution. Another drawback is that the direct linear solution does not minimize the errors in the original point and line measurements (so-called *reprojection error*) but some other quantities (so-called *algebraic error*).

These drawbacks can be prevented by determining a valid trifocal tensor by minimizing reprojection error. The so-called *Gauss-Helmert model* provides a general environment for such constrained adjustment tasks. The determination of a valid trifocal tensor in the Gauß-Helmert

model can be done basically in two ways: By parameterizing the tensor using its 27 elements and introducing the 8 internal constraints together with the scale fixing, or by using an alternative parameterization for the trifocal tensor that has 18 degrees of freedom.

Several sets of constraints and alternative parameterizations have been proposed over the years and the most important ones are reviewed in this thesis. Most of these constraints and parameterizations are derived from the so-called *tensorial slices*, which are 3×3 matrices that can be sliced out of the tensor. Due to their great importance these tensorial slices are investigated thoroughly in this thesis. Using these tensorial slices we derive two new sets of constraints together with a simple geometric interpretation, and also a new alternative parameterization for the trifocal tensor.

For computing a valid trifocal tensor by minimizing reprojection error in the Gauß-Helmert model it is important to have a consistent representation for the trifocal tensor. The parameterization using the projection matrices proposed by Hartley is the simplest way for such a consistent representation. This parameterization is applicable as long as (i) not all three projection centers coincide and (ii) the first projection center is different from the other two.

If all image features arise from a common plane in space, the trifocal tensor can not be computed uniquely. Therefore the *minimum thickness* of the object points used to compute the trifocal tensor is investigated; i.e. what is the minimum deviation of the object points from a common plane, so that the computation is still possible? This investigation will be done empirically for different image configurations and different numbers of corresponding points. In these investigations we further consider the effect on the computed tensor if the internal constraints are considered or neglected, and if algebraic error or reprojection error is minimized.

The findings of these empirical investigations can be summarized in the following way: For image configurations with strong geometry, like in the case of convergent terrestrial images or in the case of aerial images, the direct linear solution (algebraic error without constraints) for the trifocal tensor and the valid tensor from the Gauß-Helmert model (reprojection error with constraints) are practically the same. Concerning the minimum thickness unexpected small values were found already for 10 point correspondences. If the number of points increases, the minimum thickness gets even smaller. For the mentioned configurations and 15 point correspondences the computation of the trifocal tensor is *still successful* for a minimum thickness of the object points of about 1% of the camera distance (normal angle camera with assumed noise in the images of 1 pixel).

For image configurations with weak geometry, like collinear projection centers with congruent viewing directions, much more point correspondences are required and the direct linear solution fails more often for small thicknesses of the object points than the valid solution from the Gauß-Helmert model.

From these empirical investigations we come to the conclusion that minimizing reprojection error and considering the internal constraints is actually not necessary if the image configuration has a strong geometry – especially if one is only interested in initial values for a subsequent bundle block adjustment. However, if a general tool for providing such initial values is aimed for, which does not have any restrictions on the image geometry and also handles weak configurations, then the Gauß-Helmert model minimizing reprojection error must be applied.

Kurzfassung

Das Thema dieser Arbeit ist der **Trifokal-Tensor**, der die relative Orientierung (oder Epipolargeometrie) von *drei unkalibrierten Bildern* beschreibt. In diesem Sinne ist der Tensor eine Erweiterung der Fundamental-Matrix, welche die relative Orientierung von zwei unkalibrierten Bildern beschreibt. Der Trifokal-Tensor ist ein homogener Tensor der Stufe 3; dem gemäß kann er als $3\times3\times3$ Zahlenwürfel dargestellt werden. Aufgrund der folgenden Eigenschaften ist dieser Tensor von besonderem Interesse:

– Der Trifokal-Tensor kann in *linearer Weise* aus gegebenen Punkt- und Linienkorrespondenzen in drei Bildern bestimmt werden; letztere sind für die Bestimmung der Fundamental-Matrix nicht verwendbar. Aus diesem Grund stellt der Trifokal-Tensor ein *Werkzeug für die Bestimmung der relativen Orientierung von drei Bildern dar, für das keine Näherungswerte benötigt werden*. Diese so gefundene Orientierung kann dann als Startwert für eine anschließende Bündelblockausgleichung verwendet werden. Weiters ist die Wahrscheinlichkeit auf eine gefährliche Konfiguration zu treffen, die keine eindeutige Bestimmung des Tensors erlaubt, wesentlich geringer als für zwei Bilder und ihre Fundamental-Matrix. Die einzige praktisch relevante gefährliche Situation entsteht nur dann, wenn alle korrespondierenden Bildpunkte und -linien von einer gemeinsamen Ebene stammen.

  Nachdem der Trifokal-Tensor bestimmt wurde, können die Basisvektoren und Rotationsmatrizen der relativen Orientierung der drei Bilder einfach extrahiert werden – wenn die innere Orientierung bekannt ist. Ist diese unbekannt, aber ident für alle drei Bilder, so kann diese gemeinsame innere Orientierung ebenfalls im Allgemeinen bestimmt werden.

  Sind $n > 3$ Bilder gegeben, so werden mehrere unzusammenhängende Trifokal-Tensoren berechnet. Bevor alle n Bilder gemeinsam weiterverarbeitet werden können, müssen die einzelnen unzusammenhängenden Basisvektoren und Rotationsmatrizen in ein gemeinsames System transformiert werden.

– Die sogenannten *Transferbeziehungen*, die sich aus dem Trifokal-Tensor ableiten lassen, können verwendet werden, um den Inhalt (Punkte und Geraden) von zwei Quellbildern in ein anderes Bild – das Zielbild – zu transferieren. Falls dieses Zielbild eine echtes Bild ist, dann können die transferierten Positionen verwendet werden um ein Matching-Verfahren zu starten. Falls dieses Zielbild eine virtuelles Bild mit gegebener Orientierung darstellt, dann kann der Inhalt dieses neuen Bildes mit Hilfe der Transferrelationen Pixel für Pixel aus den beiden Quellbildern aufgebaut werden. Dieser Vorgang wird im Englischen als *novel view synthesis* oder *image formation* bezeichnet.

In dieser Arbeit wird der erste Punkt – die Berechnung des Trifokal-Tensors – näher behandelt.

Jedes Triple von korrespondierenden Punkten liefert 4 Gleichungen, die sogenannten *Trilinearitäten*, welche linear in den 27 Tensorelementen sind. Jedes Tripel von korrespondierenden Geraden liefert 2 lineare Gleichungen. Demzufolge benötigt man mindestens 7 Punkte, 13 Geraden oder eine passende Kombination um den Trifokal-Tensor direkt bestimmen zu können.

Mit dieser direkten linearen Lösung für den Trifokal-Tensor sind allerdings ein paar Nachteile verbunden. Der Trifokal-Tensor besteht zwar aus 27 Elementen, jedoch besitzt er nur 18 Freiheitsgrade. Aus diesem Grund müssen die Tensorelemente 8 interne Bedingungen erfüllen, neben der Festlegung des Tensormaßstabes, um einen gültigen Trifokal-Tensor zu repräsentieren.

Diese Bedingungen werden im Allgemeinen von der direkten linearen Lösung nicht erfüllt. Ein weiterer Nachteil ist, dass die direkte lineare Lösung nicht die Fehler in den originalen Bildbeobachtungen minimiert (die sogenannten *Residuen*), sondern den sogenannten *algebraischen Fehler*.

Diese Nachteile kann man beseitigen, wenn man einen gültigen Trifokal-Tensor über die Minimierung der Bildresiduen berechnet. Das sogenannte Gauß-Helmert Modell stellt eine allgemeine Umgebung für derartige bedingte Ausgleichungsaufgaben dar. Die Berechnung eines gültigen Trifokal-Tensor im Gauß-Helmert Modell kann in zwei Arten realisiert werden: Indem der Trifokal-Tensor durch seine 27 Elemente repräsentiert wird und die 8 internen Bedingungen gemeinsam mit der Festlegung des Maßstabes zusätzlich ins Gauß-Helmert Modell aufgenommen werden; oder indem der Trifokal-Tensor durch eine andere alternative Parametrisierung, die genau 18 Freiheitsgrade hat, dargestellt wird.

Verschiedene Gruppen von Bedingungen und unterschiedliche Parametrisierungen wurden in der Vergangenheit publiziert, und die wichtigsten davon werden in dieser Arbeit zusammengefasst. Die meisten dieser Bedingungen und Parametrisierungen leiten sich aus den sogenannten *Tensorschnitten* ab. Dabei handelt es sich um 3×3 Matrizen, die aus dem Tensor sozusagen herausgeschnitten werden können. Aufgrund ihrer hohen Bedeutung werden diese Tensorschnitte in dieser Arbeit sehr genau untersucht. Mit ihrer Hilfe werden zwei neue Gruppen von Bedingungen, die auch eine einfache geometrische Interpretation erlauben, und eine neue alternative Parametrisierung für den Trifokal-Tensor hergeleitet.

Um nun einen gültigen Trifokal-Tensor im Gauß-Helmert Modell über die Minimierung der Bildresiduen berechnen zu können, ist es wichtig eine konsistente Repräsentierung für den Tensor zu besitzen. Die von Hartley vorgeschlagene Parametrisierung mit Hilfe der Projektionsmatrizen ist die einfachste Möglichkeit für solch eine konsistente Repräsentierung. Diese Parametrisierung ist anwendbar solange (i) nicht alle drei Projektionszentren zusammenfallen und (ii) das erste Projektionszentrum verschieden von den anderen beiden ist.

Wenn alle Korrespondenzen in den drei Bildern von einer gemeinsamen Ebene stammen, dann kann der Trifokal-Tensor nicht eindeutig bestimmt werden. Aus diesem Grund wird in dieser Arbeit die *minimaler Dicke* der Objektpunkte untersucht; d.h. was ist die minimal notwendige Abweichung von einer gemeinsamen Ebene, sodass der Trifokal-Tensor immer noch erfolgreich berechnet werden kann? Diese Untersuchung wird empirisch anhand verschiedener Bildkonfigurationen und unterschiedlicher Punktanzahl durchgeführt. In diesen Untersuchungen werden weiters die Unterschiede im berechneten Trifokal-Tensor behandelt, die entstehen wenn die internen Bedingungen berücksichtigt werden oder nicht, und wenn algebraische Fehler oder Bildresiduen minimiert werden.

Die Erkenntnisse dieser empirischen Untersuchungen können wie folgt zusammengefasst werden: Für Bildkonfigurationen mit guter Geometrie, wie im Fall von konvergenten terrestrischen Aufnahmen oder im Fall von Luftbildern, stimmt die direkte lineare Lösung (algebraischer Fehler ohne Bedingungen) für den Tensor praktisch mit dem gültigen Tensor, der im Gauß-Helmert Modell (Bildresiduen mit Bedingungen) geschätzt wird überein. Für diese Konfigurationen mit guter Geometrie wurden auch erstaunlich geringe minimale Dicken bereits bei der Verwendung von 10 korrespondierenden Punkten gefunden. Nimmt die Anzahl der Punkte zu, so wird auch die minimale Dicke kleiner. Für die erwähnten Konfigurationen und 15 Punkttripel war die Berechung des Tensors für eine minimale Dicke von etwa 1% der Aufnahmeentfernung immer noch erfolgreich (Normalwinkel Aufnahmen und angenommenes Bildrauschen von 1 Pixel).

Für Bildkonfigurationen mit schwacher Geometrie, wie im Fall von kollinearen Projektionszentren mit zusammenfallenden Blickrichtungen, ist eine deutlich größere Punkteanzahl ist notwendig und die direkte lineare Lösung versagt viel häufiger bei kleinen Objektdicken als die gültige Lösung übers Gauß-Helmert Modell.

Diese empirischen Untersuchungen führen uns zum Schluss, dass für Konfigurationen mit guter Geometrie die Minimierung der Bildresiduen und die Berücksichtigung der internen Bedingungen eigentlich nicht notwendig ist – besonders dann nicht, wenn man nur an Näherungswerten für eine anschließende Bündelblockausgleichung interessiert ist. Benötigt man jedoch ein allgemeines Werkzeug, das Näherungswerte für die Bildorientierungen liefert und keine Einschränkungen an die Bildkonfiguration setzt und auch bei schwächerer Geometrie funktioniert, so ist die Lösung übers Gauß-Helmert Modell zu realisieren.

Contents

Chapter 1

Introduction

1.1 Overview

The topic of this thesis is the **trifocal tensor**, which describes the epipolar geometry (or relative orientation) of three uncalibrated images, just like the fundamental matrix does for two. The trifocal tensor is of particular interest because of the following properties:

– The trifocal tensor can be determined *linearly* from corresponding points *and* lines in three images. Compared to the separate linear determination of the respective three fundamental matrices from the same data, which however can only use corresponding points, the linear determination of the trifocal tensor should be more stable, as it can use also lines, and all image correspondences are used at once. Furthermore, the chances of running into critical configurations for determining the trifocal tensor is much smaller than for the fundamental matrix. The only practical relevant critical configuration for computing the trifocal tensor happens if all corresponding image features arise from a common plane in space.

  After determining the trifocal tensor, the basis vectors and rotation matrices of the relative orientation of the three images can be extracted easily if the interior orientation of the images is known. If the interior orientation is unknown, but the same for all three images, then, in general, this common interior orientation can be retrieved also.

– The *transfer relations* associated with the trifocal tensor can be used to transfer the content (points and lines) of two source images into the other one (the target image). If the target image is a *real* image, the transferred position of features can be used to initialize a matching technique in the target image. If the target image is *virtual* with a given orientation, then the transfer relations can be used to form the content of this new image by transferring pixel by pixel. This process is called *novel view synthesis* or *image formation*.

  A similar transfer can also be realized with fundamental matrices, but there the transfer of lines as entities is not possible and the transfer of points can fail for certain configurations, like collinear projection centers.

So there are many benefits in considering three images with their trifocal tensor compared to pairs of images with their fundamental matrices. In this thesis we will concentrate on the first issue of determining the trifocal tensor. Although in principle the trifocal tensor can be determined linearly, there are some drawbacks associated with this direct linear solution.

– A *valid* trifocal tensor satisfies certain *internal constraints*. The tensor obtained by the direct linear solution does in general not satisfy these internal constraints and so does not represent a valid trifocal tensor. Furthermore, depending on the number of point and line correspondences, their distribution in space and the arrangement of the three images, chances are that the redundant unconstrained elements of the tensor absorb some of the errors in the image measurements and thereby return a severely disturbed tensor and thus a rather wrong relative orientation of the three images.

– This direct linear solution does not minimize the errors in the original measurements (so-called *reprojection error*) but some other quantities (so-called *algebraic error*). And therefore we suspect that additional errors in the tensor elements are induced.

These drawbacks can be removed by determining a valid trifocal tensor by minimizing reprojection error. The so-called *Gauss-Helmert model* provides a general environment for such constrained adjustment tasks. For realizing the determination of the trifocal tensor in this model we have to investigate the internal constraints of the trifocal tensor and find ways of introducing them in this model. This can be done either by considering the necessary number of constraints explicitly or by using an alternative parameterization, which returns a valid trifocal tensor in any case.

Some set of constraints and alternative parameterizations have been presented over the years. We will review the most important ones and derive also new sets of constraints and a new parameterization. As it turns out, the simplest way of realizing a valid trifocal tensor in the Gauß-Helmert model is the parameterization using the projection matrices proposed in [Hartley 1994a], which historically is one of the first parameterizations.

If all image features arise from a common plane in space, the trifocal tensor can not be computed uniquely. Therefore it is interesting to investigate the minimum deviation from a common plane required to successfully determine the trifocal tensor. This investigation will be done empirically for different image configurations and different numbers of corresponding points. In these investigations we further consider the effect on the computed tensor if the internal constraints are considered or neglected, and if algebraic error or reprojection error is minimized.

1.2 Motivation

In this thesis concepts of two vision sciences *photogrammetry* and *computer vision* are applied. The approaches taken in these two disciplines are guided by different interests and therefore have different benefits and drawbacks. Consequently it is useful to outline the viewpoint of each science individually.

1.2.1 The photogrammetric viewpoint

The main task of photogrammetry is the reconstruction of spatial objects using a set of images. For the reconstruction the object is abstracted by a set of points and line-elements, out of which lines and surfaces are formed. The 3D coordinates of these special object points and the line-

and surface-belongings (the so-called *topology*) are to be determined. These coordinates have to be related to a – in most cases – given global coordinate system.

In a geometric manner an image can be interpreted as the result of a central projection. In photogrammetry the conventional mathematical representation of the central projection is the following; see e.g. [Kraus 1997]

$$\begin{aligned} x &= x_{p_0} - c\,\frac{r_{11}(X - X_Z) + r_{21}(Y - Y_Z) + r_{31}(Z - Z_Z)}{r_{13}(X - X_Z) + r_{23}(Y - Y_Z) + r_{33}(Z - Z_Z)} \\ y &= y_{p_0} - c\,\frac{r_{12}(X - X_Z) + r_{22}(Y - Y_Z) + r_{32}(Z - Z_Z)}{r_{13}(X - X_Z) + r_{23}(Y - Y_Z) + r_{33}(Z - Z_Z)} \end{aligned} \tag{1.1}$$

where X, Y, Z are the Cartesian coordinates of an object point. They may be known (control point) or unknown (tie point). The Cartesian coordinates x, y of the corresponding image point are observed. The principal distance c together with the coordinates of the principal point x_{p_0}, y_{p_0} form the interior orientation and are known by calibration in most cases. The coordinates of the projection center X_Z, Y_Z, Z_Z and the elements r_{ij} of a spatial rotation matrix $\boldsymbol{R}$ (from the image- to the object-coordinate-system), constructed by three rotation parameters, make up the the exterior orientation and are unknown in most cases. Altogether the representation (1.1) is done using nine parameters in a *non-linear* way.

A camera whose way of projecting can be described in the presented way is usually called a *pinhole camera model*; see also section 5.1. Photographs made by real cameras deviate more or less from an exact central projection due to errors of the lenses, film deformations or unflatness of the CCD-sensor; i.e. the pinhole model is not sufficient to describe the reality. These deviations are compensated by additional parameters including among others the *radial distortion* as the most important one. They are introduced by additive terms in the equations (1.1).

The basic requirement for doing object reconstruction with a set of images is image orientation; i.e. the estimation of the exterior and maybe the interior orientation of all the images. For some of these tasks the reference to a global system of coordinates (*absolute orientation*) is not necessary or this is done later: in other words, no or insufficient control features are available. In such cases one works with the so-called *relative orientation*. This is the alignment of at least two images in such a way that corresponding projection rays intersect in a point in space. The underlying (local) coordinate system is fixed for example in the following manner: its origin coincides with the projection center of a suitable image and the orientation of the axes of the coordinate system of that image fix the orientation of the axes of the object system.

Whether one works with a global coordinate system or with a local one of relative orientation, the equations (1.1) always give the fundamental relation between object point and image point.

Least-squares adjustment The observed image coordinates are subject to accidental errors. For decreasing their disturbing influence on the estimated unknowns the number of observations is much higher than the number of unknowns. This results in an overdetermined system, which is then solved by minimizing or maximizing a certain *criterion*. Any of the following criteria is useful. The unknown parameters are estimated such that: (i) they are unbiased and have least variance, the so-called *best unbiased estimation*, (ii) the sum of the squared residuals is minimized, the so-called *least-squares-estimation*, or (iii) the probability for the given realization of the observations is maximized, the so-called *maximum-likelihood-estimation*.

The *optimization* of the chosen criterion is realized in a certain *mathematical model*, which consists of a *stochastical model* and a *functional model*. The stochastical model describes the stochastic properties of the observations; e.g. by their covariances or their probability density functions. The functional model describes the functional relations between the observations and the unknown parameters.

Two kinds of functional models are commonly used in geodesy: the *Gauß-Markoff model*, in which each observation can be expressed in terms of the known and unknown parameters, and the *Gauß-Helmert model*, in which only functions of the observations can be expressed in terms of the parameters; see [Koch 1999] and also section D. If the observations are normally distributed, then the unknown parameters which are estimated in these two models satisfy all three criteria stated above.

In general, one works with a least-squares-estimation or *least-squares-adjustment*. But it is important to point out that such a least-squares-adjustment is a best unbiased estimation only if the sum of the squares of the errors of the original observations is minimized (so-called *reprojection* or *residual error*) and not the sum of squares of some other quantities, which are functionally dependent on groups of observations (so-called *algebraic error*).

A least-squares-adjustment based on the Gauß-Markoff model is also termed *adjustment by indirect observations*, and if it is based on the Gauß-Helmert model, it is called *general case of least squares adjustment*. The latter also allows the inclusion of constraints between the unknown parameters. We see from equation (1.1) that the observed image coordinates can be directly expressed functionally in the orientation parameters and the object coordinates. Therefore the least-squares-adjustment of images based on the pinhole camera model can be done in the Gauß-Markoff model, which in this case is also called *bundle block adjustment*.

In the course of such a bundle block adjustment the most plausible values for the unknowns (object points, the exterior and under certain circumstances also the interior orientation of the images) and their accuracies resp. co-variances are computed. But for such an adjustment linear equations are needed. Since the equations of the central projection are non-linear, they have to be linearized and the solution needs to be found *iteratively*. For this purpose approximate values of the exterior (and under certain circumstances also for the interior orientation) of the images are necessary. But in many cases the determination of these approximate values is quite tedious.

1.2.2 The computer vision viewpoint

In computer vision central perspective images are also worked with and the pinhole camera model is used, too. The tasks in computer vision that deal specially with the problem of image orientation are again object reconstruction (called *structure from motion*) and further the inspection of dynamical processes (called *3d-dynamic-scene-analysis*, *motion from structure* or *active vision*); e.g. inspections in robotics applications or tracking of moving objects like cars.

Due to the highly non-linear character of equation (1.1) in computer vision the central perspective relation between object and image is not described with the 9 elements of the interior and exterior orientation and relation (1.1), but a linear representation is aimed for. This which is especially necessary for the real time applications in active vision. This linear representation for the central projection is achieved using *projective geometry* and is obtained in form of a 3×4 matrix P – the so-called *point projection matrix*; see section 5.2. Projective geometry is very suited

for representing the geometry of vision problems, because all basic geometric entities – points, lines and planes – can simply be represented as vectors, the vector translation can be represented as a linear mapping and entities at infinity are naturally incorporated in the framework. So also the mapping of 3D lines to image lines, which is rather hard to represent in the classical notation of interior and exterior orientation, can be represented linearly using a 3×6 matrix Q – the so-called *line projection matrix*, which is closely related to P; see section 5.3.

The benefits of using projective geometry go even further, as it also allows a linear representation for the relative orientation of two, three or four images[1] by means of homogenous *multiple view tensors*. The two bifocal tensors – *essential matrix* resp. *fundamental matrix* – represent linearly the relative orientation of two calibrated resp. uncalibrated images. The *trifocal tensor* is the linear formulation of the relative orientation of three uncalibrated images and the *quadrifocal tensor* the one of four uncalibrated images. *Linear* in this case means that the relations representing the relative orientation of two, three and four images (i.e. the condition that the back projection of the corresponding image points and lines intersect in space) are linear in the respective tensor elements and multi-linear in the coordinates of the corresponding image features.

This linearity, however, is achieved for the price that these projective representations use more elements than the corresponding conventional representation with the interior and exterior orientation of the images. The homogenous multiple view tensor for n views is made up of 3^n elements but has only $11n-15$ independent elements (also called *degrees of freedom*). As a consequence $3^n-(11n-15)$ constraints, which are generally non-linear, must hold between the projective parameters of the linear formulation to represent a *valid* relative orientation of two, three or four images; see table 1.1. As long as these tensors are computed from known interior and exterior orientations of given images, this is not a problem, since then the constraints are fulfilled anyway.

number of images	number of constraints
2	2
3	9
4	16

Table 1.1: *Number of constraints for the linear formulation of the relative orientation of two, three and four uncalibrated images. Due to the homogeneity of these multiple view tensors, one of these constraints is the fixing of the scale.*

Estimating the multiple view tensors The fulfillment of the constraints constitutes a problem if such a tensor is to be determined for given image correspondences (points and/or lines) which are subject to accidental errors. The great benefit of the linearity, which makes the whole matter *independent on approximate values*, is only applicable if these constraints are neglected and algebraic error is minimized. This is the total opposite to an optimal solution, which considers the constraints and minimizes reprojection error by using the Gauß-Helmert model. The Gauss-Markoff model is not applicable in this case because the above mentioned relations for the intersections of the back projections of the corresponding image features are multi-linear functions in the image coordinates.

[1] In section C it is shown, that for more than four images no such linear representation is possible.

If the interior orientation of the images is known a-priori (especially for photogrammetric cameras), then another drawback may arise when determining these tensors linearly. This knowledge can not be used because it imposes additional non-linear constraints among the tensor elements. This has the affect that each image is individually calibrated during the linear computation. The quality of such self calibrations depends strongly on the image configuration and the distribution of the object points and lines. For example the calibration and consequently the linear determination of the tensors fails if all image correspondences arise from planar object points.

On the other hand if the interior orientation of the images is absolutely unknown (e.g. for amateur cameras, or for very old images or parts of images of unknown source) and the distribution of the points and lines permits the calibration, this feature turns into a strong benefit.

Another drawback of these projective representations is that non-linear distortion is naturally not included in this framework. This follows from the fact, that projective geometry is based on mappings called *collineations*, which map straight lines to straight lines. In the presence of non-linear distortion, however, the mapped line is no longer straight. Some strategies concerning the inclusion of non-linear distortion parameters into the projective based framework, which is not considered in this thesis, can be found in the literature; e.g. [Niini 2000] or [Fitzgibbon 2001].

1.3 Objective of the thesis

The optimal solution for the multiple view tensors requires the consideration of the actual degrees of freedom. In the Gauß-Helmert model this can be established in two ways: (i) considering the necessary number of independent constraints or (ii) using an alternative parameterization for the multiple view tensors, which has the required degrees of freedom. If this parameterization is *minimal*, i.e. the number of parameters is the same as the degrees of freedom, then no constraints need to be considered during the computation in the Gauß-Helmert model.

For the essential and fundamental matrix the constraints were investigated many years ago, e.g. [Rinner 1963], [Huang and Faugeras 1989] or [Hartley 1992], and are summarized in sections 6.1.3 and 6.1.2. Minimal parameterizations are presented in [Luong and Faugeras 1996] and [Hartley and Zisserman 2001]. Also for the quadrifocal tensor constraints and parameterizations are reported in the literature; see [Hartley 1998] and [Shashua and Wolf 2000].

The subject of this work are constraints and parameterizations for the trifocal tensor. The tensor is made up of 27 elements and has 18 degrees of freedom, thus has to satisfy 9 constraints. Several constraints (e.g. [Stein and Shashua 1998], [Papadopoulo and Faugeras 1998] ,[Canterakis 2000]) and parameterizations ([Hartley 1994a], [Torr and Zisserman 1997], [Papadopoulo and Faugeras 1998]) are reported in the literature so far. In most cases their usage in the Gauß-Helmert model, however, is limited due to one or more of the following reasons: They are

– *not complete* (< 9), like the set of [Stein and Shashua 1998]. Consequently no valid trifocal tensor will result from the computation.

– *not minimal* (> 9), like the set of [Papadopoulo and Faugeras 1998]. Therefore a suitable minimal subset of those constraints must be select for the Gauß-Helmert model, and it is not clear on which criteria this selection should be based on.

– *not valid for all image coordinate systems*, like the set of [Papadopoulo and Faugeras 1998]. Depending on the image coordinate system it may happen that certain constraints are not applicable. Although this problem can be solved by adapting the coordinate system, this requires additional unnecessary attention during the computation.

– *rather complicated to be incorporated in the Gauß-Helmert model*, like the set of [Canterakis 2000] and the minimal parameterizations of [Torr and Zisserman 1997] and [Papadopoulo and Faugeras 1998]. For this limitation we have to consider that the constraints or the relations involving the alternative parameterization must be linearized to be included in the Gauß-Helmert model. Therefore and if we want to avoid numerical differentiation, the constraints and parameterization should not be too complex.

The only parameterization so far that can be realized in the Gauß-Helmert model without difficulties is the non-minimal parameterization using the projection matrices proposed in [Hartley 1994a].

The objective of this work is to investigate the constraints of the trifocal tensor and to derive a minimal set of constraints, that can be included in the Gauß-Helmert model by explicit differentiation. Two new minimal sets of constraints presented in section 7.6.3 are quite simple compared to the only minimal set so far presented in the literature [Canterakis 2000]. They are, however, not general applicable for all image coordinate systems – just as the set of [Papadopoulo and Faugeras 1998].

We also investigate the ways the trifocal tensor can be parameterized and in section 8.3.4 we derive a new non-minimal parameterization from which a minimal parameterization can be obtained quite easily. Although this minimal parameterization is simpler than the other minimal parameterizations of [Torr and Zisserman 1997] and [Papadopoulo and Faugeras 1998], its implementation in the Gauß-Helmert model is still more complex than for the parameterization using the projection matrices.

If all image features arise from a common plane in space, the trifocal tensor can not be computed uniquely. Therefore it is interesting to investigate the minimum deviation from a common plane required to successfully determine the trifocal tensor. This investigation will be done empirically for different image configurations and different numbers of corresponding points. In these investigations we further consider the effect on the computed tensor if the internal constraints are considered or neglected, and if algebraic error or reprojection error is minimized.

1.4 Overview of the thesis

In the chapters 3 to 6 the basics, required to understand the properties of the trifocal tensor, are presented in detail.

The basics of projective geometry are presented in chapter 3, followed by a short introduction to tensor notation in chapter 4. Contrary to what we might expect from the name *trifocal tensor* the usage of tensorial algebra is rather small throughout this work, therefore this introduction is kept short.

The single-view geometry is explained in chapter 5, including the pinhole camera model, the point and line projection matrices, and the back projection of points and lines in the image to projection rays and projection planes in space.

The two-view geometry is depicted in chapter 6, including the explanation of the relative orientation of images, the epipolar geometry and the fundamental and essential matrix. This chapter is closed by the description of homographies between two images induced by a 3D plane, and correlations between two images induced by a 3D line. These two mappings between two images will appear frequently when dealing with the trifocal tensor.

The three-view geometry follows in chapter 7, which is also the main chapter of this thesis. First a derivation of the trifocal tensor is given, followed by an investigation of the tensor properties. These are derived by considering the homographies and correlations associated with the trifocal tensor; section 7.2. The *trilinearities* and *transfer relations*, which are the basic relations involving the trifocal tensor, are described in section 7.3. The internal constraints are explored in section 7.6 by reviewing existing sets and deriving new sets of constraints.

The computation of a valid trifocal tensor by minimizing reprojection error is the topic of chapter 8. After reviewing existing ways of parameterizing a valid trifocal tensor a new parameterization for the trifocal tensor is presented in section 8.3.4. In section 8.3.5 all possible ways of realizing a valid trifocal tensor in the Gauß-Helmert model are discussed critically, showing that the simplest way is the parameterization using the projection matrices.

The results of empirical investigations showing the differences in determining the trifocal tensor by minimizing algebraic error or reprojection error with or without considering the internal constraints are presented in chapter 9. Here also the effect of almost planar object points is considered.

A summary and a conclusion given in chapter 10 close this thesis.

Chapter 2

A few Words about Notation

In a rigorous way a distinction should be made between a geometric object – like a point, a line or a plane – and its mathematical representation, which is usually a vector. However, for the sake of simplicity and since all geometric objects, except for points, are of interest only in their projective representation, this distinction is not performed within this thesis and all geometric objects are addressed via their projective representation (a homogenous vector) in bold upright font. If the Euclidian representation of a point is considered, then this is indicated in bold *slanted* font.

Since we will deal with points and lines in 2D and 3D and further planes in 3D it is helpful to recognize these objects directly by their symbolical representations. Therefore the following notation was chosen, for a full summary on the used notation see section A on page 173:

2D objects:

points	
represented as Euclidian vectors:	lower case Roman font – *slanted*, e.g. $\boldsymbol{e}$, $\boldsymbol{p}$, $\boldsymbol{x}$
represented as homogenous vectors:	lower case Roman font – upright, e.g. $\mathbf{e}$, $\mathbf{p}$, $\mathbf{x}$
lines	
represented as homogenous vectors:	lower case Greek font, e.g. $\boldsymbol{\alpha}$, $\boldsymbol{\lambda}$, $\boldsymbol{\rho}$

3D objects:

points	
represented as Euclidian vectors:	upper case Roman font – *slanted*, e.g. $\boldsymbol{P}$, $\boldsymbol{X}$, $\boldsymbol{Z}$
represented as homogenous vectors:	upper case Roman font – upright, e.g. $\mathbf{P}$, $\mathbf{X}$, $\mathbf{Z}$
lines	
represented as homogenous vectors:	upper case Calligraphic font, e.g. $\boldsymbol{\mathcal{L}}$, $\boldsymbol{\mathcal{R}}$
planes	
represented as homogenous vectors:	upper case Greek font, e.g. $\boldsymbol{\Omega}$, $\boldsymbol{\Phi}$

Matrices:

Euclidian matrices (i.e. with fixed scale):	upper case Sans Serif font – slanted, e.g. R
homogenous matrices:	upper case Sans serif font – upright, e.g. C, P, I
identity matrix of dimension n:	I_n
its column vectors:	$\boldsymbol{i}_1$, $\boldsymbol{i}_2$, $\boldsymbol{i}_3$ for $n = 3$ and $\boldsymbol{I}_1$, $\boldsymbol{I}_2$, $\boldsymbol{I}_3$, $\boldsymbol{I}_4$ for $n = 4$
zero matrix of dimension n:	$\mathsf{0}_n$

Chapter 3

Projective Geometry

In this chapter the basic principles of projective geometry as far as they are of concern to this thesis are presented. Primarily this chapter is based on the introduction to projective geometry found in [Hartley and Zisserman 2001].

3.1 Introduction

As it is outlined in [Hartley and Zisserman 2001], geometry can be studied in two ways: an *approach based on geometric primitives* and an *algebraic approach*. In the first one no coordinates at all are considered and all results are based on theorems and axioms and all proofs are carried out only in terms of the geometric primitives like points, lines or planes. In the algebraic approach the geometric primitives are associated with coordinates and algebraic entities; e.g. a point is identified by a coordinate vector relative to some chosen coordinate basis. The advantage of the second approach lies in its easy amenability, due to our familiarity with coordinates and vectors since school days, and the easy possibility to derive algorithms and practical computational methods based on the results of this approach. Therefore we consider this algebraic approach.

Practically speaking, the main advantage of using projective geometry is the possibility to deal with points at infinity and to represent the translation of vectors as a linear transformation[1]. This will allow the expression of the central perspective relation between object points and corresponding image points by a single matrix multiplication.

Projective geometry is not limited to the planar case only, but can be studied in any dimension even with other fields of numbers than the real numbers $\mathbb{R}$. Since for this thesis the planar case is the most important, which is also readily accessible since it can be easily visualized, we will start with this case and then give a few notes for the three and n-dimensional case.

[1] A transformation $\sigma(\boldsymbol{x})$ is termed *linear*, if it holds: (i) $\sigma(\lambda \cdot \boldsymbol{x}) = \lambda \cdot \sigma(\boldsymbol{x})$ (for any scalar λ) and (ii) $\sigma(\boldsymbol{x} + \boldsymbol{y}) = \sigma(\boldsymbol{x}) + \sigma(\boldsymbol{y})$.

3.2 The projective plane $\mathbb{P}^2$

In the Euclidian 2D space $\mathbb{R}^2$ a point $\mathbf{x}$ is identified by a 2D column vector $\boldsymbol{x} = (x, y)^\top$[2]. A 2D line $\boldsymbol{\lambda}$ passing through the point $\mathbf{x}$ can be written as $ax + by + c = 0$. So we can represent this line by the column vector $\boldsymbol{\lambda} = (a, b, c)^\top$. The components of $\boldsymbol{\lambda}$ have the following meanings: $\boldsymbol{n}_\lambda = (a, b)^\top$ is the normal vector of the line and $|c|/|\boldsymbol{n}_\lambda|$ is the perpendicular distance of the line to the origin, where $c = -\boldsymbol{n}_\lambda^\top \boldsymbol{x}$.

It is easy to see that the very same line can be represented also by the relation $(ka)x + (kb)y + (kc) = 0$ and thus by the vector $k(a, b, c)^\top$ for any non-zero scalar k. These two vectors, which differ only by their scaling, are considered as equivalent. Vectors with this property are called **homogenous vectors**. Consequently, although $\boldsymbol{\lambda}$ has 3 elements it only has 2 **degrees of freedom**. Considering this homogenity we actually may not write $\boldsymbol{\lambda} = (a, b, c)^\top$, because the symbol '=' is misplaced, and thus in the following we will write

$$\boldsymbol{\lambda} \sim \begin{pmatrix} a \\ b \\ \hline c \end{pmatrix} \sim \begin{pmatrix} \boldsymbol{\lambda}_H \\ \lambda_O \end{pmatrix} \tag{3.1}$$

whenever homogenous vectors are involved; '$\sim$' emphasizes that the vector on the left and/or right side can be multiplied with any non-singular scalar anytime. Nevertheless, we will use '=' sometimes, to emphasize that the scale of a projective quantity is initialized by a specific relation. The separation of a projective vector $\mathbf{p}$ in a homogenous part '$\mathrm{p_H}$', which does not depend on the distance to the origin, and a Euclidian part '$\mathrm{p_O}$', which does, was suggested by [Brand 1966] and will be helpful in some of the following applications. Both parts can appear as scalars (being p_H or p_O) or as vectors (being $\boldsymbol{p}_H$ or $\boldsymbol{p}_O$).

Now, if we address the 2D point $\mathbf{x}$ with the vector $\mathbf{x}$

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \rightarrow \quad \mathbf{x} \sim \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} u \\ v \\ \hline w \end{pmatrix} \sim \begin{pmatrix} \boldsymbol{x}_O \\ x_H \end{pmatrix} \tag{3.2}$$

we see that the **incidence relation** $ax + by + c = 0$, that $\mathbf{x}$ lies on $\boldsymbol{\lambda}$, can be expressed as

$$\boldsymbol{\lambda}^\top \mathbf{x} = 0. \tag{3.3}$$

And we see, that $\mathbf{x}$ is also a homogenous vector with 2 degrees of freedom, since it holds $a(kx) + b(ky) + c(k{\cdot}1) = 0$. This leads to the following (simplified) definition of the **projective space** $\mathbb{P}^2$ (in other words the *projective plane*):

Definition 3.1 (The projective plane.) *The projective space $\mathbb{P}^2$ consists of a set of points $\mathfrak{P}$, a set of lines $\mathfrak{L}$ and an incidence relation* (3.3)*, that defines whether a point $\mathbf{x}$ lies on a line $\boldsymbol{\lambda}$. Each of both sets is represented by the set of homogenous vectors ($\mathbf{x}$ resp. $\boldsymbol{\lambda}$) of $\mathbb{R}^3 \backslash \{\mathbf{0}\}$. The zero vector $\mathbf{0}$ must be excluded from $\mathbb{R}^3$ since it does not correspond to any valid point or line.*

[2]Rigorously $\boldsymbol{x} = (x, y)^\top$ just represents the *vector of components* or *vector of coordinates* with respect to some arbitrarily chosen vector basis $\widetilde{\boldsymbol{B}} = [\widetilde{\boldsymbol{b}}_1, \widetilde{\boldsymbol{b}}_2]$, where $\widetilde{\ }$ symbolizes a *vector in space* – $\mathbb{R}^2$ in this case. The vector $\widetilde{\boldsymbol{x}}$ in space corresponding to the vector of components $\boldsymbol{x}$ is obtained by $\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{B}}\boldsymbol{x}$. However, if the vector basis is chosen to be the *canonical basis* (i.e. $\widetilde{\boldsymbol{b}}_1 = (1, 0)^\top$ and $\widetilde{\boldsymbol{b}}_2 = (0, 1)^\top$), then the vector of components $\boldsymbol{x}$ will be the same as the vector in space $\widetilde{\boldsymbol{x}}$ and thus allowing a simpler notation. So in this thesis always the canonical basis will be assumed for defining a particular space.

If $\mathbf{x} \sim (u, v, w)^\top \sim \left(\boldsymbol{x}_O^\top, x_H\right)^\top$ is a general homogenous or **projective vector** of a 2D point $\mathbf{x}$, then the corresponding vector $\boldsymbol{x}$ in $\mathbb{R}^2$ is obtained by (provided $w \neq 0$ resp. $x_H \neq 0$):

$$\mathbf{x} \sim \begin{pmatrix} u \\ v \\ w \end{pmatrix} \sim \begin{pmatrix} \boldsymbol{x}_O \\ x_H \end{pmatrix} \quad \rightarrow \quad \boldsymbol{x} = \begin{pmatrix} u/w \\ v/w \end{pmatrix} = \frac{1}{x_H} \boldsymbol{x}_O \tag{3.4}$$

Since points $\mathbf{x}$ and lines $\boldsymbol{\lambda}$ of $\mathbb{P}^2$ are represented as 3D vectors, $\mathbb{P}^2$ can be visualized according to figure 3.1. Thus projective points $\mathbf{x}$ of $\mathbb{P}^2$ can be interpreted as *3D lines* passing through $\mathbf{0}$ and projective lines $\boldsymbol{\lambda}$ of $\mathbb{P}^2$ can be interpreted as *3D planes* going through $\mathbf{0}$. The vector $\boldsymbol{\lambda}$ interpreted as a vector in $\mathbb{R}^3$ is the normal vector of this plane. The intersection of these lines and planes with the $w = 1$ plane gives the Euclidian points and lines.

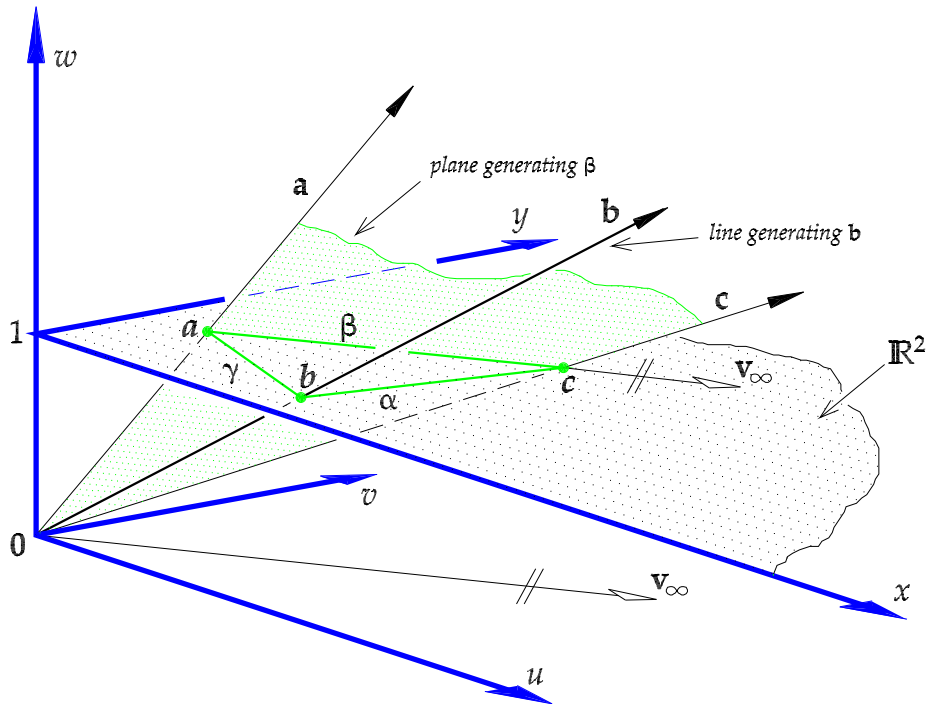

Figure 3.1: *The projective space $\mathbb{P}^2$ interpreted in $\mathbb{R}^3$, and the Euclidian space $\mathbb{R}^2$.*

This figure also indicates the problem if the $w-$coordinate of $\mathbf{x}$ is 0. In this case no valid intersection point with the $w = 1$ plane exists. Such points (like $\mathbf{v}_\infty$ in the figure) are called **points at infinity**. They lie infinitely far away and represent *directions of 2D lines*. Since all points at infinity are represented as homogenous 3D vectors with $w = 0$, the set of all points at infinity is 1-dimensional giving a line $\boldsymbol{\pi}_\infty \sim (0, 0, 1)^\top$, which is called **line at infinity**. This gives another interpretation for the projective plane: $\mathbb{P}^2$ is obtained by $\mathbb{R}^2$ and adding the line at infinity.

If we consider the incidence relation (3.3) for points and lines again and fix the line $\boldsymbol{\lambda}$, then (3.3) is a *point-wise* representation for the fixed line $\boldsymbol{\lambda}$. Points lying on a single line are called **collinear points**. A set of collinear points is also called a **pencil of points**. Conversely we can fix the point $\mathbf{x}$ and consider all possible lines which satisfy (3.3) and thus get a *line-wise* representation for the fixed point $\mathbf{x}$. Lines intersecting in one single point are called **concurrent lines**. A set of concurrent lines is also called a **pencil of lines**.

Duality. These last considerations and the fact that points and lines in $\mathbb{P}^2$ are both represented by 3D vectors, thus given an arbitrary 3D vector, one can not determine whether it is a point or a line without further information, indicate that these entities are dual to each other.

This **duality** of points and lines in $\mathbb{P}^2$ suggests the possibility to swap the set of lines $\mathfrak{L}$ with the set of points $\mathfrak{P}$. Doing this one obtains the so-called *dual projective plane* $\mathbb{P}^{2*}$ made up of $\mathfrak{P}^*$ and $\mathfrak{L}^*$. So when dealing with $\mathbb{P}^{2*}$ the lines of $\mathbb{R}^2$ are identified as "points" in $\mathbb{P}^{2*}$ (i.e. the set $\mathfrak{P}^*$) and the points of $\mathbb{R}^2$ are identified as "lines" in $\mathbb{P}^{2*}$ (i.e. the set $\mathfrak{L}^*$). The major gain of considering the dual projective plane is that any theorem valid in $\mathbb{P}^2$ is also valid in $\mathbb{P}^{2*}$. This has the consequence that the dual of the theorem is also valid in $\mathbb{P}^2$. This is the so-called **duality principle**. The dual of a theorem is obtained by interchanging the words "points" and "lines", by keeping the incidences and by interchanging the words "intersect" and "join"[3]. For example the two basic 2D configurations – a set of concurrent lines and a set of collinear points – are dual to each other.

Theorem of Desargues. In section 7.6.3 we will use the following theorem, see also figure 3.2.

Theorem 3.1 (Theorem of Desargues.) *Let two triangles in* $\mathbb{P}^2$ *be defined by the points* $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ *and* $\{\bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{c}}\}$. *The lines* $\mathbf{a}\bar{\mathbf{a}}$, $\mathbf{b}\bar{\mathbf{b}}$, $\mathbf{c}\bar{\mathbf{c}}$ *intersect in a single point* $\mathbf{z}$ *if and only if the intersections of corresponding sides* ($\mathbf{ab}$, $\bar{\mathbf{a}}\bar{\mathbf{b}}$), ($\mathbf{bc}$, $\bar{\mathbf{b}}\bar{\mathbf{c}}$), ($\mathbf{ca}$, $\bar{\mathbf{c}}\bar{\mathbf{a}}$) *lie on a single line* $\boldsymbol{\lambda}$.

Equivalently, we can say if two triangles are **perspective from the point z**, then they are **perspective from a line λ**.

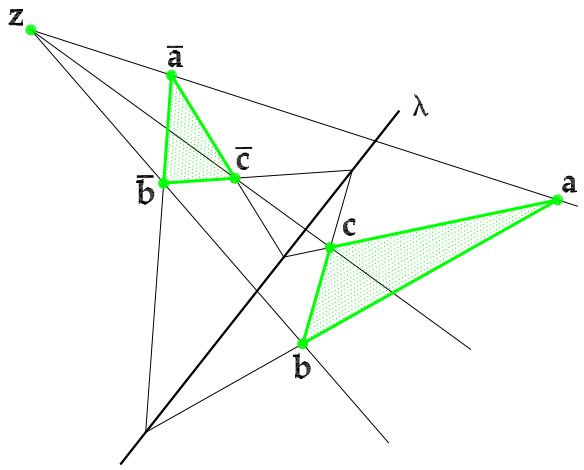

Figure 3.2: *The theorem of Desargues.*

Interpretation of the canonical basis vectors of $\mathbb{R}^3$. The columns of I_3 represent the canonical basis vectors of $\mathbb{R}^3$ and can be interpreted as the homogenous representations of points or lines in $\mathbb{P}^2$; see table 3.1 and figure 3.3.

[3]The terms *join* and *intersect* are defined in section 3.5.

column vector	interpretation as point	interpretation as line
$\boldsymbol{i}_1 = (1,0,0)^\top$	point at infinity of the x-axis	y-axis
$\boldsymbol{i}_2 = (0,1,0)^\top$	point at infinity of the y-axis	x-axis
$\boldsymbol{i}_3 = (0,0,1)^\top$	origin of the coordinate system	line at infinity $\boldsymbol{\pi}_\infty$

Table 3.1: *Interpretation of the columns of $\boldsymbol{I}_3$ as points and lines in $\mathbb{P}^2$.*

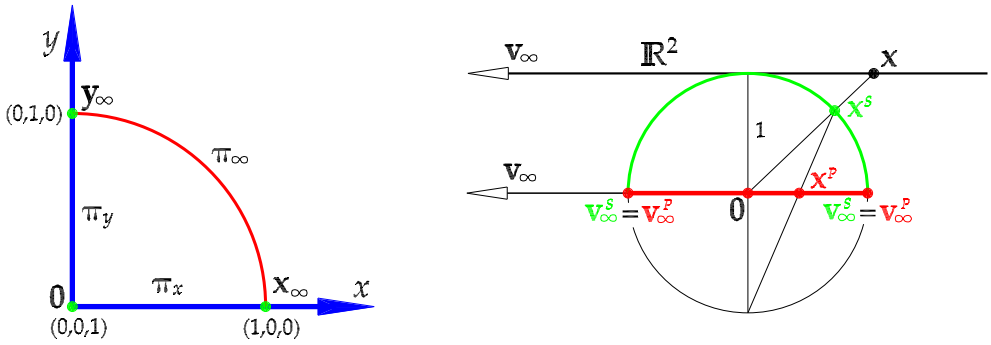

Figure 3.3: *The interpretation of the columns of $\boldsymbol{I}_3$ as points and lines in $\mathbb{P}^2$ is shown in the left part. Points: Origin $\mathbf{0}$, point at infinity of the x-axis $\mathbf{x}_\infty$, point at infinity of the y-axis $\mathbf{y}_\infty$. Lines: x-axis $\boldsymbol{\pi}_x$, y-axis $\boldsymbol{\pi}_y$, line at infinity $\boldsymbol{\pi}_\infty$. Note, that $\boldsymbol{\pi}_\infty$ is drawn as a circle. This map of $\mathbb{P}^2$ is created by the stereographic projection shown in the right part: First all projective points $\mathbf{x}$ are normalized to $|\mathbf{x}| = 1$, therefore the resulting points $\mathbf{x}^S$ lie all on the north hemisphere of the unit sphere. This way the points at infinity are mapped to the points on the equator of that sphere; e.g. $\mathbf{v}_\infty$. Afterwards all these points on the unit sphere are projected from the south pole onto the equator plane resulting in the points $\mathbf{x}^P$. Consequently the map of $\boldsymbol{\pi}_\infty$ becomes a circle with radius 1, and no other points lie outside this circle. In the left part of the figure the super-script P is omitted.*

Conics. A **conic** is a curve of degree two in $\mathbb{P}^2$. The points $\mathbf{x} \sim (u,v,w)^\top$, which lie on a conic, satisfy the following equation

$$au^2 + buv + cuw + dv^2 + evw + fw^2 = 0.$$

Using the matrix C this can be written also as

$$\mathbf{x}^\top \mathsf{C} \mathbf{x} = 0, \tag{3.5}$$

with

$$\mathsf{C} \sim \begin{bmatrix} a & b/2 & c/2 \\ b/2 & d/2 & e/2 \\ c/2 & e/2 & f \end{bmatrix}.$$

The matrix C is a representation of a conic in $\mathbb{P}^2$; C is homogenous and symmetric and therefore has 5 degrees of freedom. If the matrix C is non-singular the conic is termed **non-degenerate**, and degenerate otherwise. The conic (3.5) was defined using points. Due to the duality principle we can also define a **dual conic** C^* by its tangent lines $\boldsymbol{\lambda}$; i.e.

$$\boldsymbol{\lambda}^\top \mathsf{C}^* \boldsymbol{\lambda} = 0. \tag{3.6}$$

The matrices of a non-degenerate conic and its dual are related by $\mathsf{C}^* = \mathsf{C}^{-1}$.

In Euclidean geometry a non-degenerate conic can be one of the following types: circle, ellipse, parabola or hyperbola. A degenerate conic is either a pair of lines ($\text{rank}(\mathsf{C}) = 2$) or one repeated line ($\text{rank}(\mathsf{C}) = 1$).

3.3 The projective spaces $\mathbb{P}^3$ and $\mathbb{P}^n$

Basically $\mathbb{P}^3$ is just a generalization of $\mathbb{P}^2$. So, a 3D point $\mathbf{X}$ can be addressed with a projective vector $\mathbf{X}$

$$\boldsymbol{X} = (X, Y, Z)^\top \quad \rightarrow \quad \mathbf{X} \sim \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \sim \left(\begin{array}{c} P \\ Q \\ R \\ \hline S \end{array}\right) \sim \begin{pmatrix} \boldsymbol{X}_O \\ X_H \end{pmatrix}. \tag{3.7}$$

The projective representation of 3D lines is not that straight forward and will follow at the end of this section. A *3D plane*, however, can be represented very easily. If a 3D plane $\boldsymbol{\Omega}$ is given by the equation $aX + bY + cZ + d = 0$, then the homogenous column vector

$$\boldsymbol{\Omega} = \left(\begin{array}{c} a \\ b \\ c \\ \hline d \end{array}\right) \sim \begin{pmatrix} \boldsymbol{\Omega}_H \\ \Omega_O \end{pmatrix} \tag{3.8}$$

is the projective representation of this 3D plane. The components of $\boldsymbol{\Omega}$ have the following meaning: $\boldsymbol{\Omega}_H = (a, b, c)^\top$ is the normal vector $\boldsymbol{n}_\Omega$ of the plane and $|d|/|\boldsymbol{n}_\Omega| = |\Omega_O|/|\boldsymbol{\Omega}_H|$ is the perpendicular distance of the plane to the origin. The *incidence relation* $aX + bY + cZ + d = 0$, that $\mathbf{X}$ lies on $\boldsymbol{\Omega}$ can be expressed as

$$\boldsymbol{\Omega}^\top \mathbf{X} = 0, \tag{3.9}$$

and so $\Omega_O = -\boldsymbol{\Omega}_H^\top \boldsymbol{X}$.

Since both, 3D points and 3D planes, are represented by homogenous vectors of $\mathbb{R}^4$, a duality between points and planes exists in $\mathbb{P}^3$ – whereas in $\mathbb{P}^2$ a duality between points and lines could be observed. The general statement is that in $\mathbb{P}^n$ subspaces of dimension k are dual to subspaces of dimension $n - k - 1$. As a consequence dual subspaces span the whole $\mathbb{P}^n$, if they are skew (i.e. their intersection is $\{\}$). So in $\mathbb{P}^n$ dual to points (having dimension 0) are so-called **hyperplanes** (having dimension $n - 1$). Thus the join of a hyperplane and a point outside span the whole $\mathbb{P}^n$. The hyperplanes are lines in $\mathbb{P}^2$ and planes in $\mathbb{P}^3$. The basic 3D configurations – a set of concurrent planes (called **bundle of planes**[4]) and a set of coplanar points – are dual in $\mathbb{P}^3$.

If $\mathbf{X} \sim (P, Q, R, S)^\top \sim \left(\boldsymbol{X}_O^\top, X_H\right)^\top$ is a general homogenous vector of a 3D point $\mathbf{X}$, then, provided $S \neq 0$ resp. $X_H \neq 0$, the corresponding vector $\boldsymbol{X}$ in $\mathbb{R}^3$ is obtained by:

$$\mathbf{X} \sim \begin{pmatrix} P \\ Q \\ R \\ S \end{pmatrix} \quad \rightarrow \quad \boldsymbol{X} = \begin{pmatrix} P/S \\ Q/S \\ R/S \end{pmatrix} = \frac{1}{X_H} \boldsymbol{X}_O \tag{3.10}$$

[4] A *pencil of planes* is made up of planes sharing a common 3D line. A *pencil* is a 1-parameter family, whereas a *bundle* is a 2-parameter family.

And again if $S = X_H = 0$, then $\mathbf{X}$ represents a point at infinity which is not part of $\mathbb{R}^3$. All 3D points at infinity lie in a plane – the so-called **plane at infinity**, represented by the projective vector $\mathbf{\Pi}_\infty = (0, 0, 0, 1)^\top$.

Interpretation of the canonical basis vectors of $\mathbb{R}^4$. The columns of I_4 represent the canonical basis vectors of $\mathbb{R}^4$ and can be interpreted as the homogenous representations of points or planes in $\mathbb{P}^3$; see table 3.2 and figure 3.4.

column vector	interpretation as point	interpretation as plane
$\boldsymbol{I}_1 = (1, 0, 0, 0)^\top$	point at infinity of the X-axis	YZ-plane
$\boldsymbol{I}_2 = (0, 1, 0, 0)^\top$	point at infinity of the Y-axis	ZX-plane
$\boldsymbol{I}_3 = (0, 0, 1, 0)^\top$	point at infinity of the Z-axis	XY-plane
$\boldsymbol{I}_4 = (0, 0, 0, 1)^\top$	origin of the coordinate system	plane at infinity $\mathbf{\Pi}_\infty$

Table 3.2: *Interpretation of the columns of I_4 as points and planes in $\mathbb{P}^3$.*

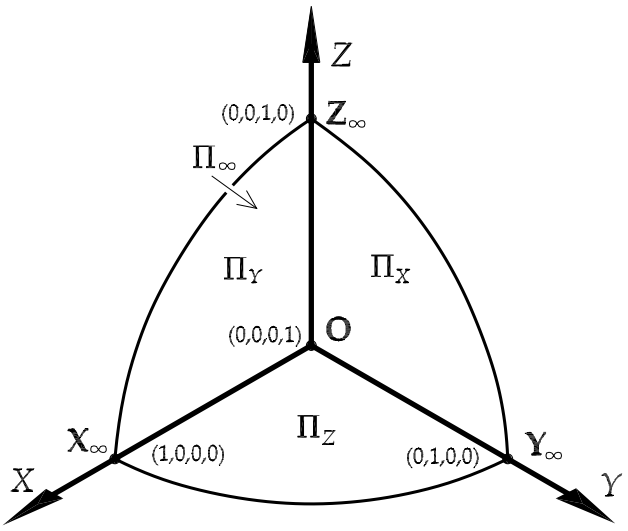

Figure 3.4: *Interpretation of the columns of I_4 as points and planes in $\mathbb{P}^3$. Points: Origin $\mathbf{0}$, point at infinity of X-axis $\mathbf{X}_\infty$, point at infinity of Y-axis $\mathbf{Y}_\infty$, point at infinity of Z-axis $\mathbf{Z}_\infty$. Planes: YZ-plane $\mathbf{\Pi}_X$, XZ-plane $\mathbf{\Pi}_Y$, XY-plane $\mathbf{\Pi}_Z$, plane at infinity $\mathbf{\Pi}_\infty$. Note, that the lines at infinity are drawn as circles. This map of $\mathbb{P}^3$ is created by the stereographic projection;cf. text of figure 3.3 on page 15.*

3.3.1 The Plücker line coordinates

Lines in 3D space have 4 degrees of freedom. This can be seen easily by intersecting a given 3D line $\boldsymbol{\mathcal{L}}$ with two planes; e.g. the $X = 0$ plane and the $Z = 0$ plane. The resulting two intersection points define $\boldsymbol{\mathcal{L}}$ and have 4 degrees of freedom. For that reason a 3D line $\boldsymbol{\mathcal{L}}$ can not be represented as a homogenous vector in $\mathbb{P}^3$. This explains, why 3D lines require a special

treatment compared to 3D points and 3D planes which can be represented in $\mathbb{P}^3$ (as homogenous vectors of $\mathbb{R}^4$). However, as it will be shown, 3D lines can be represented as homogenous vectors in $\mathbb{P}^5$.

The basic idea for the 3D line representation is due to Julius Plücker, a German geometrician of the 19^{th} century. He proposed a representation based on $\binom{n+1}{k+1}$ coordinates for any k-dimensional subspace within an n-dimensional projective space. These so-called Plücker coordinates arise as the $(k+1) \times (k+1)$ minors of an $(n+1) \times (k+1)$ matrix M, whose columns are made up of the k defining points of the k-dimensional subspace. Of the $\binom{n+1}{k+1}$ Plücker coordinates only $(k+1)(n-k)$ are independent; cf. [Faugeras and Luong 2001].

Following Plücker's idea we can compute a coordinate representation for 3D lines. In this case we have $k = 1$ and $n = 3$. A 3D line $\boldsymbol{\mathcal{L}}$ is defined by two 3D points $\mathbf{A}$ and $\mathbf{B}$, which are represented by $\mathbf{A} = (A_1, A_2, A_3, A_4)^\top$ and $\mathbf{B} = (B_1, B_2, B_3, B_4)^\top$. $\mathbf{A}$ and $\mathbf{B}$ define the columns of a matrix M as follows:

$$\mathsf{M} = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \\ A_4 & B_4 \end{bmatrix}.$$

Then we have to compute the $(k+1) \times (k+1)$ (i.e. 2×2) minors of M. These minors can be used to build up another matrix, called *Plücker matrix*, $\mathsf{L}(\mathbf{A}, \mathbf{B})$ as

$$L_{ij} = A_i B_j - B_i A_j,$$

or in other words:

$$\mathsf{L} = \mathbf{A}\mathbf{B}^\top - \mathbf{B}\mathbf{A}^\top. \tag{3.11}$$

Since $L_{ij} = -L_{ji}$ and $L_{ii} = 0$, L is a 4×4 skew-symmetric matrix and only 6 elements of L are essential, which is exactly the number of coordinates proposed by Plücker (i.e. $\binom{4}{2}$). There are several possibilities to choose the six essential elements. The following are commonly chosen for good reason and make up the **six homogenous Plücker coordinates of a line $\boldsymbol{\mathcal{L}}$** represented by $\boldsymbol{\mathcal{L}}$

$$\boldsymbol{\mathcal{L}} \sim \begin{pmatrix} \boldsymbol{L}_H \\ \boldsymbol{L}_O \end{pmatrix} \sim (L_{41}, L_{42}, L_{43}, \quad L_{23}, L_{31}, L_{12})^\top. \tag{3.12}$$

The vector $\boldsymbol{\mathcal{L}}$ is homogenous (since $\mathsf{L}(\alpha\mathbf{A}, \beta\mathbf{B}) = \alpha\beta\mathsf{L}(\mathbf{A}, \mathbf{B})$) and $\boldsymbol{\mathcal{L}} \neq \mathbf{0}$ if the defining points are different. $\boldsymbol{\mathcal{L}}$ is also independent from the choice of the defining points, which can be verified readily by showing $\mathsf{L}(\mathbf{C}, \mathbf{D}) = (\alpha\delta - \beta\gamma)\mathsf{L}(\mathbf{A}, \mathbf{B})$ for $\mathbf{C} = \alpha\mathbf{A} + \beta\mathbf{B}$ and $\mathbf{D} = \gamma\mathbf{A} + \delta\mathbf{B}$.

Since a 3D line has only 4 degrees of freedom, but is represented as a homogenous vector of dimension 6, one additional constraint must hold between the 6 elements of $\boldsymbol{\mathcal{L}}$, which is the so-called **Plücker-identity**:

$$\boldsymbol{L}_H^\top \boldsymbol{L}_O = 0. \tag{3.13}$$

A projective geometric interpretation for (3.13) is provided by the so-called *Klein model*, which states, that lines in $\mathbb{R}^3$ are mapped to points in $\mathbb{P}^5$. But not any point in $\mathbb{P}^5$ may represent a 3D line, only those points in $\mathbb{P}^5$ which lie on the quadratic surface defined by (3.13), the so-called *Klein quadric*, correspond to a line in $\mathbb{R}^3$.

A Euclidian geometric interpretation for the components $\boldsymbol{L}_H$ and $\boldsymbol{L}_O$ of the Plücker coordinate vector $\boldsymbol{\mathcal{L}}$ is obtained by considering the homogenous and Euclidian parts of the defining points

A and **B**; i.e. $\mathbf{A} \sim (\boldsymbol{A}_O^\top, A_H)^\top$ and $\mathbf{B} \sim (\boldsymbol{B}_O^\top, B_H)^\top$. It can be easily verified, that in this case $\boldsymbol{L}_H$ and $\boldsymbol{L}_O$ follow as

$$\boldsymbol{\mathcal{L}} \sim \begin{pmatrix} \boldsymbol{L}_H \\ \boldsymbol{L}_O \end{pmatrix} \sim \begin{pmatrix} A_H \boldsymbol{B}_O - B_H \boldsymbol{A}_O \\ \boldsymbol{A}_O \times \boldsymbol{B}_O \end{pmatrix}. \tag{3.14}$$

By considering $A_H = B_H = 1$ we see, that $\boldsymbol{L}_H$ represents the direction of the line and $\boldsymbol{L}_O$ represents the normal vector of the plane joining $\boldsymbol{\mathcal{L}}$ and the origin; cf. figure 3.5. This simple geometric interpretation justifies our choice of essential parameters in (3.12).

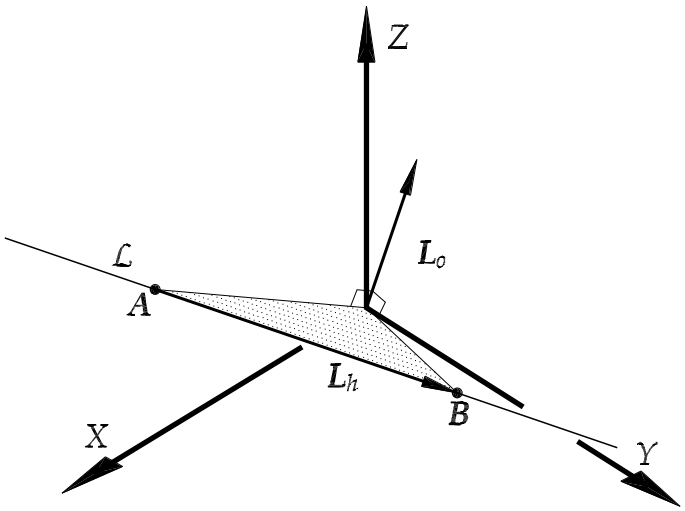

Figure 3.5: *The vectors $\boldsymbol{L}_H$ and $\boldsymbol{L}_O$, which make up the Plücker coordinates of a 3D line $\boldsymbol{\mathcal{L}}$.*

So we see that relation (3.14) provides an easy way of computing the Plücker coordinate vector $\boldsymbol{\mathcal{L}}$ for a line $\boldsymbol{\mathcal{L}}$ joining two real points **A** and **B**. If $\boldsymbol{\mathcal{L}}$ is defined by one real point $\boldsymbol{A}$ and the direction vector **d**, the Plücker coordinate vector $\boldsymbol{\mathcal{L}}$ can be computed as

$$\boldsymbol{\mathcal{L}} \sim \begin{pmatrix} \boldsymbol{L}_H \\ \boldsymbol{L}_O \end{pmatrix} \sim \begin{pmatrix} A_H \mathbf{d} \\ \boldsymbol{A}_O \times \mathbf{d} \end{pmatrix}. \tag{3.15}$$

Using the Plücker coordinates of 3D lines it can be checked easily if two lines $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{M}}$ intersect. Suppose $\boldsymbol{\mathcal{L}}$ is defined by the points **A** and **B** and the line $\boldsymbol{\mathcal{M}}$ is defined by the points **C** and **D**, then the condition for $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{M}}$ to intersect is identical to the condition that the four defining points lie in one common plane (they are **coplanar**). In other words the four points are linearly dependent and the matrix $[\mathbf{ABCD}]$ has to have zero determinant:

$$\begin{aligned}
&|\mathbf{ABCD}| = 0 \\
&= \begin{vmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A_3 & B_3 & C_3 & D_3 \\ A_4 & B_4 & C_4 & D_4 \end{vmatrix} = \begin{vmatrix} A_1 & B_1 \\ A_4 & B_4 \end{vmatrix} \begin{vmatrix} C_2 & D_2 \\ C_3 & D_3 \end{vmatrix} - \begin{vmatrix} A_2 & B_2 \\ A_4 & B_4 \end{vmatrix} \begin{vmatrix} C_1 & D_1 \\ C_3 & D_3 \end{vmatrix} + \begin{vmatrix} A_3 & B_3 \\ A_4 & B_4 \end{vmatrix} \begin{vmatrix} C_1 & D_1 \\ C_2 & D_2 \end{vmatrix} \\
&\quad + \begin{vmatrix} A_2 & B_2 \\ A_3 & B_3 \end{vmatrix} \begin{vmatrix} C_1 & D_1 \\ C_4 & D_4 \end{vmatrix} - \begin{vmatrix} A_1 & B_1 \\ A_3 & B_3 \end{vmatrix} \begin{vmatrix} C_2 & D_2 \\ C_4 & D_4 \end{vmatrix} + \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \begin{vmatrix} C_3 & D_3 \\ C_4 & D_4 \end{vmatrix} \\
&\quad = +L_{14}M_{23} - L_{24}M_{13} + L_{34}M_{12} + L_{23}M_{14} - L_{13}M_{24} + L_{12}M_{34} \\
&\quad = -L_{41}M_{23} - L_{42}M_{31} - L_{43}M_{12} - L_{23}M_{41} - L_{31}M_{42} - L_{12}M_{43}.
\end{aligned}$$

In the development of $|\mathbf{ABCD}|$ we find the Plücker coordinates of $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{M}}$ and so can write the condition $|\mathbf{ABCD}| = 0$ as

$$-\boldsymbol{L}_H^\top \boldsymbol{M}_O - \boldsymbol{L}_O^\top \boldsymbol{M}_O = 0.$$

This is the incidence relation (or **coplanarity condition**) for two 3D lines and can also be interpreted as the inner product of two 3D lines

$$<\boldsymbol{\mathcal{L}}, \boldsymbol{\mathcal{M}}> = -\boldsymbol{L}_H^\top \boldsymbol{M}_O - \boldsymbol{L}_O^\top \boldsymbol{M}_H = \boldsymbol{\mathcal{L}}^\top \boldsymbol{\mathcal{M}}^*. \tag{3.16}$$

Actually, inner products are defined for coordinates of dual spaces. This can be seen also in the equations (3.3) and (3.9) on pages 12 and 16 respectively, where the incidence relations between points and lines resp. planes are formulated as inner products. Consequently

$$\boldsymbol{\mathcal{M}}^* = D\boldsymbol{\mathcal{M}} = \begin{pmatrix} -\boldsymbol{M}_O \\ -\boldsymbol{M}_H \end{pmatrix}, \tag{3.17}$$

with

$$D = \begin{bmatrix} 0_3 & -I_3 \\ -I_3 & 0_3 \end{bmatrix}, \tag{3.18}$$

must be the dual of a 3D line $\boldsymbol{\mathcal{M}}$. Observe that $\boldsymbol{\mathcal{M}}$ and $\boldsymbol{\mathcal{M}}^*$ agree up to order[5]. Of course the dual of a 3D line could be found also by applying the principle of duality; i.e. by the intersection of two 3D planes $\boldsymbol{\Omega}$ and $\boldsymbol{\Phi}$. By considering the dual version of relation (3.11) and using the same essential elements as in (3.12) the dual 3D line $\boldsymbol{\mathcal{M}}^*$ is found to be

$$\boldsymbol{\mathcal{M}}^* \sim \begin{pmatrix} \Omega_H \boldsymbol{\Phi}_O - \Phi_H \boldsymbol{\Omega}_O \\ \boldsymbol{\Omega}_O \times \boldsymbol{\Phi}_O \end{pmatrix}, \tag{3.19}$$

with the defining planes $\boldsymbol{\Omega} \sim (\boldsymbol{\Omega}_H^\top, \Omega_O)^\top$ and $\boldsymbol{\Phi} \sim (\boldsymbol{\Phi}_H^\top, \Phi_O)^\top$.

Although Plücker coordinates may appear a bit confusing at first sight, they are very useful when working in 3D space. This will be shown in section 3.5 on page 23 where 3D lines are used to construct points and planes.

Interpretation of the canonical basis vectors of $\mathbb{R}^6$. The columns of I_6 represent the canonical basis vectors of $\mathbb{R}^6$ and can be interpreted as the homogenous representations of lines in $\mathbb{P}^3$; see table 3.3 and figure 3.6.

3.4 Projective transformations in $\mathbb{P}^n$

Before we discuss projective transformations in $\mathbb{P}^n$, a few remarks need to be given concerning the definition of a basis in $\mathbb{P}^n$. Since $\mathbb{P}^n$ is based on $\mathbb{R}^{n+1}$, $n+1$ points are needed to define a vector basis $\overline{\mathfrak{B}} = \{\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_{n+1}\}$ for $\mathbb{R}^{n+1}$. With respect to this basis $\overline{\mathfrak{B}}$ a projective vector $\mathbf{X}$ can be assigned components $\{X_1, X_2, ..., X_{n+1}\}$; i.e. $\mathbf{X} \sim \sum_i X_i \mathbf{B}_i$. Since the basis vectors $\mathbf{B}_i$ are homogenous vectors, they can be replaced by $s_i \mathbf{B}_i$ anytime ($s_i \neq 0$), thus the components of $\mathbf{X}$ would depend on the chosen scale of the $\mathbf{B}_i$. This is compensated by introducing an additional

[5]The change of sign in D does not alter the homogenous relations involving $\boldsymbol{\mathcal{L}}^*$, but aligns the definition of $\boldsymbol{\mathcal{L}}^*$ with the one found in [Heuel 2002], where the dual line is defined parallel to the so-called Grassmann-Caley algebra; cf. [Faugeras and Luong 2001].

column vector	interpretation as line
$\boldsymbol{\mathcal{I}}_1 = (1,0,0,0,0,0)^\top$	X-axis
$\boldsymbol{\mathcal{I}}_2 = (0,1,0,0,0,0)^\top$	Y-axis
$\boldsymbol{\mathcal{I}}_3 = (0,0,1,0,0,0)^\top$	Z-axis
$\boldsymbol{\mathcal{I}}_4 = (0,0,0,1,0,0)^\top$	line at infinity of the YZ-plane
$\boldsymbol{\mathcal{I}}_5 = (0,0,0,0,1,0)^\top$	line at infinity of the ZX-plane
$\boldsymbol{\mathcal{I}}_6 = (0,0,0,0,0,1)^\top$	line at infinity of the XY-plane

Table 3.3: *Interpretation of the columns of $\boldsymbol{I}_6$ as lines in $\mathbb{P}^3$.*

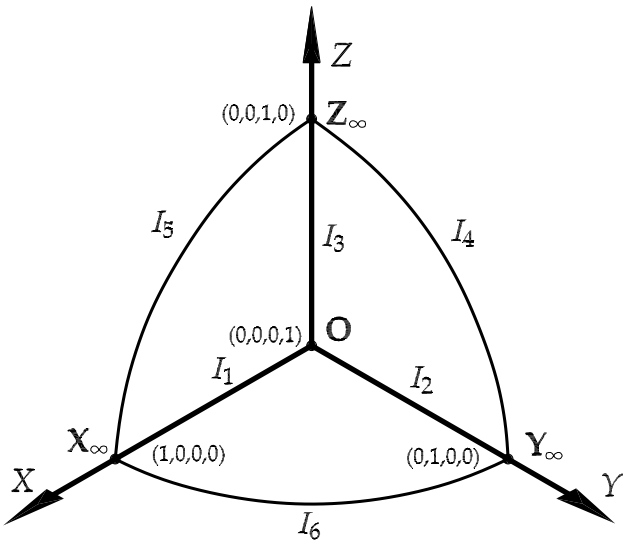

Figure 3.6: *Interpretation of the columns of $\boldsymbol{I}_6$ as lines in $\mathbb{P}^3$; cf. table 3.3. The same stereographic projection as in figure 3.4 on page 17 is applied.*

base point $\mathbf{B}_0$, the so-called *point of unity*, with the property $\mathbf{B}_0 = \sum_i \mathbf{B}_i$ which linearly depends on *all* $n+1$ basis vectors in $\overline{\mathfrak{B}}$. With this point of unity the internal scale of the $n+2$ base points is fixed, hence the name, and the *homogenous* components $\{X_1, X_2, ..., X_{n+1}\}$ of a projective vector $\mathbf{X}$ with respect to the basis $\mathfrak{B} = \{\mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_{n+1}\}$ are unique. Usually the canonical basis will be assumed. In this case the point of unity is given by the $(n+1)$-dimensional vector $(1, 1, \ldots, 1)$. These considerations result in the following theorem:

Theorem 3.2 (Basis of $\mathbb{P}^n$.) *A basis of $\mathbb{P}^n$ is made up of $n+2$ points – with no $n+1$ points being linearly dependent.*

The main projective transformation is the **collineation** or **homography** which is geometrically defined in the following way:

Definition 3.2 (Collineation.) *A mapping* h *from $\mathbb{P}^n$ onto itself is called collineation if* h *is invertible and if collinear points are mapped to collinear points.*

Algebraically this means that h can be represented by a regular $(n+1) \times (n+1)$ matrix H. Consequently, for points in $\mathbb{P}^2$ a collineation is represented as:

$$\overline{\mathbf{x}} \sim h(\mathbf{x}) \quad \rightarrow \quad \overline{\mathbf{x}} \sim \mathsf{H}\mathbf{x}. \tag{3.20}$$

Obviously, H is also a homogenous quantity and thus has $n(n+2)$ degrees of freedom; i.e. 8 in case of $\mathbb{P}^2$ and 15 in case of $\mathbb{P}^3$. Since each pair of corresponding points $\{\mathbf{x}, \overline{\mathbf{x}}\}$ gives n independent equations a collineation in $\mathbb{P}^n$ is defined by $n+2$ points $\mathbf{x}_i$ and their corresponding partners $\overline{\mathbf{x}}_i$; i.e. a projective basis is required in both systems $\mathbf{x}_i$ and $\overline{\mathbf{x}}_i$.

A special case is the **perspective collineation**, which is a collineation h that has a fixed line $\boldsymbol{\alpha}$ (a hyperplane in general), which is called *axis* whose points remain fixed under h. Due to the duality theorem a perspective collineation must also have a fixed point $\mathbf{z}$, which is called *center* of the perspective collineation. If the center lies on the axis, the perspective collineation is termed *elation* and *homology* otherwise. If no fixed element exists the collineation is termed a **projective collineation**.

Equation (3.20) describes the projective map of a point $\mathbf{x}$ in $\mathbb{P}^2$. Due to the duality principle a similar relation must be available for lines (or hyperplanes in general). If points $\mathbf{x}$ lie on a line $\boldsymbol{\lambda}$ (thus $\boldsymbol{\lambda}^\top \mathbf{x} = 0$) and are mapped according to $\overline{\mathbf{x}} \sim \mathsf{H}\mathbf{x}$, then it holds $\overline{\boldsymbol{\lambda}}^\top \overline{\mathbf{x}} = 0$ and we get

$$\boldsymbol{\lambda} \sim \mathsf{H}^\top \overline{\boldsymbol{\lambda}}, \tag{3.21}$$

and if H is regular the mapping of lines is given by:

$$\overline{\boldsymbol{\lambda}} \sim \mathsf{H}^{-\top} \boldsymbol{\lambda}. \tag{3.22}$$

In definition 3.2 on the page before a collineation was defined as a mapping of points to points. It is also possible to consider mappings from points to hyperplanes. Such transformations are called **correlations**:

Definition 3.3 *A mapping* g *from* $\mathbb{P}^n$ *onto itself is called correlation if* g *is invertible and if collinear points are mapped to concurrent hyperplanes.*

The algebraic expression is similar to (3.20), meaning that a correlation is represented by a $(n+1) \times (n+1)$ matrix G, which has $n(n+2)$ degrees of freedom. In $\mathbb{P}^2$ the correlation of points to lines is represented as

$$\overline{\boldsymbol{\lambda}} \sim g(\mathbf{x}) \quad \rightarrow \quad \overline{\boldsymbol{\lambda}} \sim \mathsf{G}\mathbf{x}. \tag{3.23}$$

The dual version of a correlation (i.e. a mapping of hyperplanes to points) will be termed **dual correlation**. We will encounter correlations in various chapters of this thesis; e.g. in form of the fundamental matrix in chapter 6.1.2. Observe that a correlation can be interpreted as a mapping of the points of $\mathbb{P}^{n*}$ to the points of $\mathbb{P}^n$ and thus the properties of collineations can be applied – especially relation (3.21).

Although collineations and correlations were associated with regular matrices, they nevertheless can be singular matrices. In this case they do not map the entire $\mathbb{P}^n$ onto itself, but only subspaces of $\mathbb{P}^n$ to subspaces of $\mathbb{P}^n$. The relations (3.20) and (3.21) hold nevertheless.

Observe that the columns of any transformation matrix H or G of dimension $n \times m$ can be interpreted as the mappings of the column vectors of the identity matrix I_m; i.e. the mappings of the respective canonical basis.

The projective transformation of 3D lines is represented by a 6×6 matrix $\mathsf{H}_{\mathcal{L}}$. Its derivation is given in section 3.5.1 on page 26 utilizing some concepts presented in the next section.

3.5 Constructing new entities in $\mathbb{P}^2$ and $\mathbb{P}^3$

In the previous sections the basic geometric entities point, line and plane together with their representation as homogenous vectors were introduced. The next important thing is to consider how these entities are to be handled algebraically to derive or *construct* new entities; e.g. to build a line by joining two points or to determine the point of intersection of a plane and a line.

At first one has to consider the possible operations that can be done in this context. Actually there are just two such operations: join and intersection. They are dual to each other and were already mentioned during the point-line-duality in section 3.2. The **join** operation takes given entities and extends them to the smallest projective subspace that contains the given entities; e.g. two points define a line. The **intersection** (or *meet*) operation does exactly the opposite: it takes the largest projective subspace which is incident to the given entities; e.g. two planes intersect in a line; [Faugeras and Luong 2001].

The result of a join or intersect operation only corresponds to a valid geometrical entity if the result is non-empty. Algebraically this means that the resulting vector $\mathbf{r}$ is different from the zero vector $\mathbf{0}$. The result nevertheless can lie at infinity. Algebraically this means that the homogenous parts $\boldsymbol{r}_H$ or r_H of the the resulting vector $\mathbf{r}$ are identical to $\mathbf{0}$ or 0. The join operation will be denoted by '$\vee$' and the intersection by '$\wedge$'.[6]

In $\mathbb{P}^2$ we only have two possible operations: The construction of a new line $\boldsymbol{\lambda}$ by the join of two points $\mathbf{x}$ and $\mathbf{y}$ and the construction of a new point $\mathbf{x}$ by the intersection of two lines $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$. The algebraic expressions of these operations are as follows:

$$\mathbf{x} \vee \mathbf{y} = \boldsymbol{\lambda} \quad \rightarrow \quad \boxed{\boldsymbol{\lambda} \sim \mathbf{x} \times \mathbf{y} = \mathsf{S}(\mathbf{x})\mathbf{y}} \tag{3.24}$$

$$\boldsymbol{\lambda} \wedge \boldsymbol{\mu} = \mathbf{x} \quad \rightarrow \quad \boxed{\mathbf{x} \sim \boldsymbol{\lambda} \times \boldsymbol{\mu} = \mathsf{S}(\boldsymbol{\lambda})\boldsymbol{\mu}.} \tag{3.25}$$

The rank-2 matrix $\mathsf{S}(\mathbf{a})$ used for the cross product $\times$ in (3.24) and (3.25) is the so-called *axiator* sometimes represented as $[\,\mathbf{a}\,]_\times$. For $\mathbf{a} \sim (u, v, w)^\top$ it is given by (see also section B.3 on page 179)

$$\mathbf{a} \times \mathbf{b} = \mathsf{S}(\mathbf{a})\mathbf{b} = -\mathsf{S}(\mathbf{b})\mathbf{a} = \mathsf{S}^\top(\mathbf{b})\mathbf{a} \quad \text{with} \quad \mathsf{S}(\mathbf{a}) = \begin{bmatrix} 0 & -w & v \\ w & 0 & -u \\ -v & u & 0 \end{bmatrix}. \tag{3.26}$$

In $\mathbb{P}^3$ we have several possibilities of joining and intersecting points, lines and planes. Following [Heuel 2002] these constructions can be expressed as simple matrix multiplications. First we consider a 3D line $\boldsymbol{\mathcal{L}}$, which is created by the **join of two points X** and **Y**; cf. relation (3.14) on page 19.

$$\boldsymbol{\mathcal{L}} = \mathbf{X} \vee \mathbf{Y} \quad \rightarrow \quad \boldsymbol{\mathcal{L}} \sim \begin{pmatrix} \boldsymbol{L}_H \\ \boldsymbol{L}_O \end{pmatrix} \sim \begin{pmatrix} X_H \boldsymbol{Y}_O - Y_H \boldsymbol{X}_O \\ \boldsymbol{X}_O \times \boldsymbol{Y}_O \end{pmatrix} = \begin{bmatrix} X_H \boldsymbol{I}_3 & -\boldsymbol{X}_O \\ \mathsf{S}(\boldsymbol{X}_O) & \mathsf{0} \end{bmatrix} \begin{pmatrix} \boldsymbol{Y}_O \\ Y_H \end{pmatrix}$$

In a short form we have

$$\boxed{\boldsymbol{\mathcal{L}} \sim \mathsf{\Pi}(\mathbf{X})\mathbf{Y}} \tag{3.27}$$

[6] The symbol $\vee$ means 'join' whereas the symbol $\cup$ means 'union' of sets. The union of two points $\mathbf{a}$ and $\boldsymbol{b}$ is just these two points, whereas the join of two points $\mathbf{a}$ and $\boldsymbol{b}$ contains all points (also called *span*) that can be expressed as linear combinations of $\mathbf{a}$ and $\mathbf{b}$.

with

$$\Pi(\mathbf{X}) := \begin{bmatrix} X_H I_3 & -\boldsymbol{X}_O \\ \mathsf{S}(\boldsymbol{X}_O) & \mathbf{0} \end{bmatrix} = \left[\begin{array}{ccc|c} X_4 & 0 & 0 & -X_1 \\ 0 & X_4 & 0 & -X_2 \\ 0 & 0 & X_4 & -X_3 \\ \hline 0 & -X_3 & X_2 & 0 \\ X_3 & 0 & -X_1 & 0 \\ -X_2 & X_1 & 0 & 0 \end{array}\right] \text{ with } \mathbf{X} = \begin{pmatrix} \boldsymbol{X}_O \\ X_H \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix}. \quad (3.28)$$

If the 3D line $\boldsymbol{\mathcal{L}}$ is to be constructed as the **intersection of two planes $\boldsymbol{\Omega}$ and $\boldsymbol{\Phi}$** then we just have to consider the dual of $\boldsymbol{\mathcal{L}} = \mathbf{X} \vee \mathbf{Y}$ which is

$$\boxed{\boldsymbol{\mathcal{L}}^* = \boldsymbol{\Omega} \wedge \boldsymbol{\Phi}}$$

and this can be represented as

$$\boldsymbol{\mathcal{L}}^* \sim \Pi(\boldsymbol{\Omega})\boldsymbol{\Phi} \quad (3.29)$$

$$\boxed{\boldsymbol{\mathcal{L}} \sim D\Pi(\boldsymbol{\Omega})\boldsymbol{\Phi} = \Pi^*(\boldsymbol{\Omega})\boldsymbol{\Phi}} \quad (3.30)$$

$$\Pi^*(\boldsymbol{\Omega}) := D\Pi(\boldsymbol{\Omega}) = \begin{bmatrix} -\mathsf{S}(\boldsymbol{\Omega}_H) & \mathbf{0} \\ -\Omega_O I_3 & \boldsymbol{\Omega}_H \end{bmatrix} \quad (3.31)$$

using D from equation (3.18) on page 20 and $\boldsymbol{\Omega} \sim (\boldsymbol{\Omega}_H^\top, \Omega_O)^\top$.

If we fix $\Pi(\boldsymbol{\Omega})$ in equation (3.29) then $\Pi(\boldsymbol{\Omega})$ describes a collineation from $\mathbb{P}^{3*}$ to $\mathbb{P}^{4*}$; i.e. from Φ to $\mathcal{L}$. Following relation (3.21) on page 22 we get the respective dual collineation as

$$\Pi^{\top}(\boldsymbol{\Omega})\boldsymbol{\mathcal{L}}^{**} \sim \boldsymbol{\Phi}^*.$$

Observing that the dual of $\boldsymbol{\mathcal{L}}^*$ is $\boldsymbol{\mathcal{L}}$ and the dual of a plane is a point, we found the relation for a point $\mathbf{X}$ as the **intersection of a line $\boldsymbol{\mathcal{L}}$ and a plane $\boldsymbol{\Omega}$**

$$\boxed{\mathbf{X} \sim \Pi^{\top}(\boldsymbol{\Omega})\boldsymbol{\mathcal{L}}.} \quad (3.32)$$

We can reformulate this last relation in the following way

$$\mathbf{X} \sim \Pi^{\top}(\boldsymbol{\Omega})\boldsymbol{\mathcal{L}} = \begin{bmatrix} \Omega_O I_3 & -\mathsf{S}(\boldsymbol{\Omega}_H) \\ -\boldsymbol{\Omega}_H^\top & \mathbf{0}^\top \end{bmatrix} \begin{pmatrix} \boldsymbol{L}_H \\ \boldsymbol{L}_O \end{pmatrix} = \begin{bmatrix} \mathsf{S}(\boldsymbol{L}_O) & \boldsymbol{L}_H \\ -\boldsymbol{L}_H^\top & 0 \end{bmatrix} \begin{pmatrix} \boldsymbol{\Omega}_H \\ \Omega_O \end{pmatrix},$$

which gives

$$\boxed{\mathbf{X} \sim \Gamma^{\top}(\boldsymbol{\mathcal{L}})\boldsymbol{\Omega}} \quad (3.33)$$

using

$$\Gamma(\boldsymbol{\mathcal{L}}) := \begin{bmatrix} -\mathsf{S}(\boldsymbol{L}_O) & -\boldsymbol{L}_H \\ \boldsymbol{L}_H^\top & 0 \end{bmatrix} = \left[\begin{array}{ccc|c} 0 & L_6 & -L_5 & -L_1 \\ -L_6 & 0 & L_4 & -L_2 \\ L_5 & -L_4 & 0 & -L_3 \\ \hline L_1 & L_2 & L_3 & 0 \end{array}\right] \text{ with } \boldsymbol{\mathcal{L}} = \begin{pmatrix} \boldsymbol{L}_H \\ \boldsymbol{L}_O \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \end{pmatrix}. \quad (3.34)$$

Note that $\Gamma(\mathcal{L})$ is identical to the Plücker matrix presented in equation (3.11) on page 18. Next we consider the construction of planes. We get a plane $\mathbf{\Omega}$ by the **join of a line $\mathcal{L}$ and a point $\mathbf{X}$** by considering the dual version of the equation (3.32)

$$\boxed{\mathbf{\Omega} \sim \overline{\Pi}(\mathbf{X})\mathcal{L}^* = \overline{\Pi^*}(\mathbf{X})\mathcal{L}} \tag{3.35}$$

or the dual of equation (3.33)

$$\boxed{\mathbf{\Omega} \sim \overline{\Gamma}(\mathcal{L}^*)\mathbf{X} = \overline{\Gamma^*}(\mathcal{L})\mathbf{X}} \tag{3.36}$$

with

$$\Gamma(\mathcal{L}^*) := \Gamma^*(\mathcal{L}) = \begin{bmatrix} \mathsf{S}(\boldsymbol{L}_H) & \boldsymbol{L}_O \\ -\boldsymbol{L}_O^\top & 0 \end{bmatrix}. \tag{3.37}$$

A plane $\mathbf{\Omega}$ can also be constructed as the **join of three points $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$**

$$\mathbf{\Omega} = \mathbf{X}_1 \vee \mathbf{X}_2 \vee \mathbf{X}_3 = (\mathbf{X}_1 \vee \mathbf{X}_2) \vee \mathbf{X}_3 = \mathcal{L} \vee \mathbf{X}_3.$$

Using (3.27) and (3.36) this can algebraically be represented as

$$\mathbf{\Omega} \sim \overline{\Gamma^*}(\Pi(\mathbf{X}_1)\mathbf{X}_2)\mathbf{X}_3. \tag{3.38}$$

Dual to this is a point $\mathbf{X}$ defined as the **intersection of three planes $\mathbf{\Omega}_1$, $\mathbf{\Omega}_2$ and $\mathbf{\Omega}_3$** which can be represented as

$$\mathbf{X} \sim \overline{\Gamma^*}(\Pi(\mathbf{\Omega}_1)\mathbf{\Omega}_2)\mathbf{\Omega}_3. \tag{3.39}$$

All the constructions presented in this section are bi- or trilinear in the participating entities and are summarized in table 3.4. The algebraic representations of these constructions return a unique result if the entities are in *general position* and an empty result otherwise. Note that the intersection of two 3D lines does not fit into this schema as this task also returns an empty result even for general (i.e. skew) lines.

entities	construction	expression
points $\mathbf{x}$,$\mathbf{y}$	$\boldsymbol{\lambda} = \mathbf{x} \vee \mathbf{y}$	$\boldsymbol{\lambda} \sim \mathsf{S}(\mathbf{x})\mathbf{y} = -\mathsf{S}(\mathbf{y})\mathbf{x}$
lines $\boldsymbol{\lambda}$,$\boldsymbol{\mu}$	$\mathbf{x} = \boldsymbol{\lambda} \wedge \boldsymbol{\mu}$	$\mathbf{x} \sim \mathsf{S}(\boldsymbol{\lambda})\boldsymbol{\mu} = -\mathsf{S}(\boldsymbol{\mu})\boldsymbol{\lambda}$
points $\mathbf{X}$, $\mathbf{Y}$	$\mathcal{L} = \mathbf{X} \vee \mathbf{Y}$	$\mathcal{L} \sim \Pi(\mathbf{X})\mathbf{Y} = -\Pi(\mathbf{Y})\mathbf{X}$
planes $\mathbf{\Omega}$, $\mathbf{\Phi}$	$\mathcal{L} = \mathbf{\Omega} \wedge \mathbf{\Phi}$	$\mathcal{L} \sim \Pi^*(\mathbf{\Omega})\mathbf{\Phi} = -\Pi^*(\mathbf{\Phi})\mathbf{\Omega}$
point $\mathbf{X}$, line $\mathcal{L}$	$\mathbf{\Omega} = \mathbf{X} \vee \mathcal{L}$	$\mathbf{\Omega} \sim \overline{\Gamma^*}(\mathcal{L})\mathbf{X} = \overline{\Pi^*}(\mathbf{X})\mathcal{L}$
plane $\mathbf{\Omega}$, line $\mathcal{L}$	$\mathbf{X} = \mathbf{\Omega} \wedge \mathcal{L}$	$\mathbf{X} \sim \overline{\Pi}(\mathbf{\Omega})\mathcal{L} = \overline{\Gamma}(\mathcal{L})\mathbf{\Omega}$
points $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$	$\mathbf{\Omega} = \mathbf{X}_1 \vee \mathbf{X}_2 \vee \mathbf{X}_3$	$\mathbf{\Omega} \sim \overline{\Gamma^*}(\Pi(\mathbf{X}_1)\mathbf{X}_2)\mathbf{X}_3$
planes $\mathbf{\Omega}_1, \mathbf{\Omega}_2, \mathbf{\Omega}_3$	$\mathbf{X} = \mathbf{\Omega}_1 \wedge \mathbf{\Omega}_2 \wedge \mathbf{\Omega}_3$	$\mathbf{X} \sim \overline{\Gamma^*}(\Pi(\mathbf{\Omega}_1)\mathbf{\Omega}_2)\mathbf{\Omega}_3$

Table 3.4: *Algebraic representations of constructions using join ($\vee$) and intersect ($\wedge$) of points, lines and planes in $\mathbb{P}^2$ and $\mathbb{P}^3$; cf. [Heuel 2002].*

Note that the relations (3.24), (3.25), (3.27), (3.30), (3.32), (3.33), (3.35) and (3.36) can be used to check the **identity** and **incidence of two entities**. For these purpose these relations have to return the zero vector, because in the case of congruence or identity the respective objects do not create a new object by join or intersection; cf. [Heuel 2002].

3.5.1 Projective transformations of lines in $\mathbb{P}^3$

If the coordinate system $\mathfrak{S}$ is changed into another one $\overline{\mathfrak{S}}$ by a (regular) collineation h, which transforms points according to

$$\overline{\mathbf{X}} \sim \mathsf{H}\mathbf{X},$$

where H is a regular homogenous 4×4 matrix, then $\mathsf{H}^{-\top}$ describes the mapping of the planes

$$\overline{\mathbf{\Omega}} \sim \mathsf{H}^{-\top}\mathbf{\Omega}.$$

The corresponding transformation of 3D lines can be represented by a regular homogenous 6×6 matrix $\mathsf{H}_{\mathcal{L}}$ as

$$\overline{\mathcal{L}} \sim \mathsf{H}_{\mathcal{L}}\mathcal{L}.$$

The following proposition shows how $\mathsf{H}_{\mathcal{L}}$ can be determined for a given H.

Proposition 3.1 (Transformation of 3D lines.) *Let the point transformation be represented by a regular homogenous* 4×4 *matrix* H*, which can be depicted as*

$$\mathsf{H} = \begin{bmatrix} \mathsf{A} & \mathbf{t} \\ \mathbf{v}^\top & w \end{bmatrix}.$$

Then the line transformation matrix $\mathsf{H}_{\mathcal{L}}$ *is given by*

$$\mathsf{H}_{\mathcal{L}} = \begin{bmatrix} w\mathsf{A} - \mathbf{t}\mathbf{v}^\top & -\mathsf{A}\mathsf{S}(\mathbf{v}) \\ \mathsf{S}(\mathbf{t})\mathsf{A} & \mathrm{adj}(\mathsf{A}) \end{bmatrix},$$

where $\mathrm{adj}(\mathsf{A})$ *is the adjoint matrix (see relation* (B.14) *on page 179). The matrix* A *is required to have* $\mathrm{rank} \geq 2$*, otherwise* H *would be singular. If* A *is regular, then* $\mathrm{adj}(\mathsf{A}) = \det(\mathsf{A})\mathsf{A}^{-1}$.

Proof:
The derivation follows the one given in [Heuel 2002], where, however, only a special case of H is considered, namely when A is a 3D rotation matrix and $\mathbf{v} = \mathbf{0}$.

The ith column of $\mathsf{H}_{\mathcal{L}}$ is the map of the ith canonical 3D line $\mathcal{I}_i$; see table 3.3 on page 21. The line $\mathcal{I}_i$ itself is the join of the jth and kth canonical point $\boldsymbol{I}_j$ and $\boldsymbol{I}_k$, of $\mathbb{P}^3$ shown in table 3.2 on page 17. The relation between i, j and k follows from the Plücker coordinates in equation (3.12) on page 18 and is given in the table below.

i	1	2	3	4	5	6
j	4	4	4	2	3	1
k	1	2	3	3	1	2

Consequently the ith column of $\mathsf{H}_{\mathcal{L}}$, denoted by $\mathbf{h}_{\mathcal{L}i}$, can be expressed by the join of the maps of the points $\boldsymbol{I}_j$ and $\boldsymbol{I}_k$; i.e. $\mathbf{h}_{\mathcal{L}i} = \mathsf{H}\boldsymbol{I}_j \vee \mathsf{H}\boldsymbol{I}_k$. Using the join-relation (3.27) on page 23 this can be expressed as $\mathbf{h}_{\mathcal{L}i} = \mathsf{\Pi}(\mathsf{H}\boldsymbol{I}_j)\mathsf{H}\boldsymbol{I}_k$.

With the structure of the operator $\mathsf{\Pi}()$ given in equation (3.28) on page 24 and the form of H in proposition 3.1 we get $\mathbf{h}_{\mathcal{L}i}$ as

$$\mathbf{h}_{\mathcal{L}i} = \begin{bmatrix} v_j\boldsymbol{I}_3 & \maltese \\ -\mathsf{S}(\maltese) & 0 \end{bmatrix} \mathsf{H}\boldsymbol{I}_k,$$

with

$$\maltese = \begin{cases} \mathsf{A}_j & \text{if j<4} \\ \mathbf{t} & \text{if j=4} \end{cases},$$

where A_j is the jth column of A and $\mathsf{S}()$ is the axiator defined in section 3.5 on page 23. By considering the six possible triples (i, j, k) and the axiator relations given in section B.3 on page 179 the form of $\mathsf{H}_{\mathcal{L}}$ is obtained after a few manipulations.

□

3.5.2 Transformation of $\mathsf{S}()$, $\mathsf{\Pi}()$ and $\mathsf{\Gamma}()$ for a given point transformation H

If the underlying coordinate system $\mathfrak{S}$ is changed into another one $\overline{\mathfrak{S}}$ by a (regular) collineation h, which transforms points, lines and planes according to

$$\begin{array}{ccc} \overline{\mathbf{x}} \sim \mathsf{H}\mathbf{x}, & \overline{\boldsymbol{\lambda}} \sim \mathsf{H}^{-\top}\boldsymbol{\lambda}, & \text{resp.} \\ \overline{\mathbf{X}} \sim \mathsf{H}\mathbf{X}, & \overline{\boldsymbol{\Omega}} \sim \mathsf{H}^{-\top}\boldsymbol{\Omega}, & \overline{\mathcal{L}} \sim \mathsf{H}_{\mathcal{L}}\mathcal{L}\,, \end{array}$$

where H is a regular homogenous 3×3 or 4×4 matrix and $\mathsf{H}_{\mathcal{L}}$ is given by proposition 3.1 on the facing page, then also the construction matrices $\mathsf{S}()$, $\mathsf{\Pi}()$ and $\mathsf{\Gamma}()$ will change in a certain way.

The change of $\mathsf{S}()$ is simple and is given by the axiator relation (B.12) on page 179

$$\mathsf{S}(\overline{\mathbf{x}}) = \mathsf{S}(\mathsf{H}\mathbf{x}) = \det(\mathsf{H})\mathsf{H}^{-\top}\mathsf{S}(\mathbf{x})\mathsf{H}^{-1} \tag{3.40}$$

$$\mathsf{S}(\overline{\boldsymbol{\lambda}}) = \mathsf{S}(\mathsf{H}^{-\top}\boldsymbol{\lambda}) = \det(\mathsf{H})\mathsf{H}\mathsf{S}(\boldsymbol{\lambda})\mathsf{H}^{\top}. \tag{3.41}$$

Observe that H is a regular homogenous 3×3 matrix, thus H^{-1} exists and $\det(\mathsf{H}) \neq 0$. As long as this factor is non-zero it is actually not important for the projective relations involving $\mathsf{S}()$.

The derivation for the transformation of $\mathsf{\Pi}()$ and $\mathsf{\Gamma}()$ is given next. With relation (3.27) on page 23 the 3D line $\mathcal{L}$ created by the join of two points $\mathbf{X}$ and $\mathbf{Y}$ can be represented as $\mathcal{L} = \mathsf{\Pi}(\mathbf{X})\mathbf{Y}$. The corresponding relation in system $\overline{\mathfrak{S}}$ is given by $\overline{\mathcal{L}} = \mathsf{\Pi}(\overline{\mathbf{X}})\overline{\mathbf{Y}}$. With $\overline{\mathcal{L}} = \mathsf{H}_{\mathcal{L}}\mathcal{L}$ and $\overline{\mathbf{Y}} = \mathsf{H}\mathbf{Y}$ we have $\mathsf{H}_{\mathcal{L}}\mathcal{L} = \mathsf{\Pi}(\overline{\mathbf{X}})\mathsf{H}\mathbf{Y}$ or $\mathsf{H}_{\mathcal{L}}\mathsf{\Pi}(\mathbf{X})\mathbf{Y} = \mathsf{\Pi}(\overline{\mathbf{X}})\mathsf{H}\mathbf{Y}$. Since this relation has to hold for any $\mathbf{Y}$ we get

$$\mathsf{\Pi}(\overline{\mathbf{X}}) = \mathsf{H}_{\mathcal{L}}\mathsf{\Pi}(\mathbf{X})\mathsf{H}^{-1}. \tag{3.42}$$

By similar considerations we find the transformation for $\mathsf{\Pi}(\boldsymbol{\Omega})$, $\mathsf{\Gamma}(\mathcal{L})$ and $\mathsf{\Gamma}(\mathcal{L}^*)$ to be

$$\mathsf{\Pi}(\overline{\boldsymbol{\Omega}}) = \mathsf{\Pi}(\mathsf{H}^{-\top}\boldsymbol{\Omega}) = {\mathsf{H}_{\mathcal{L}}}^{-\top}\mathsf{\Pi}(\boldsymbol{\Omega})\mathsf{H}^{\top} \tag{3.43}$$

$$\mathsf{\Gamma}(\overline{\mathcal{L}}) = \mathsf{\Gamma}(\mathsf{H}_{\mathcal{L}}\mathcal{L}) = \mathsf{H}\mathsf{\Gamma}(\mathcal{L})\mathsf{H}^{\top} \tag{3.44}$$

$$\mathsf{\Gamma}(\overline{\mathcal{L}}^*) = \mathsf{\Gamma}({\mathsf{H}_{\mathcal{L}}}^{-\top}\mathcal{L}^*) = \mathsf{H}^{-\top}\mathsf{\Gamma}(\mathcal{L}^*)\mathsf{H}^{-1}. \tag{3.45}$$

3.6 Checking the identity of points and lines in $\mathbb{P}^2$

If two points or two lines are identical then their join or intersection does not create a valid projective entity. The algebraic expression for *no projective entity* is the zero vector. Therefore

the axiator $\mathsf{S}()$ is a suitable means to check the identity of two projective points or lines in $\mathbb{P}^2$ by

$$\begin{aligned} \mathbf{x} \equiv \mathbf{y} \quad &\leftrightarrow \quad \mathsf{S}(\mathbf{x})\mathbf{y} = \mathbf{0}, \\ \boldsymbol{\lambda} \equiv \boldsymbol{\mu} \quad &\leftrightarrow \quad \mathsf{S}(\boldsymbol{\lambda})\boldsymbol{\mu} = \mathbf{0}. \end{aligned}$$

Observe that although $\mathbf{0}$ has 3 components, this check, however, has only 2 degrees of freedom, because points and lines have 2 independent elements in $\mathbb{P}^2$. Also, this can be seen from the rank of $\mathsf{S}()$, which is 2. Therefore we can drop one row in $\mathsf{S}()$, the one that is closest to the zero vector, and obtain a reduced matrix $\mathsf{S}^{red}()$ to perform the identity check as

$$\mathbf{x} \equiv \mathbf{y} \quad \leftrightarrow \quad \mathsf{S}^{red}(\mathbf{x})\mathbf{y} = \mathbf{0}, \tag{3.46}$$

$$\boldsymbol{\lambda} \equiv \boldsymbol{\mu} \quad \leftrightarrow \quad \mathsf{S}^{red}(\boldsymbol{\lambda})\boldsymbol{\mu} = \mathbf{0}. \tag{3.47}$$

For real points the third coordinate will always be non-zero. So we can always use the following reduced version $\mathsf{S}^{red}(\mathbf{x})$ for points, obtained by keeping the first two rows of $\mathsf{S}(\mathbf{x})$,

$$\mathsf{S}^{red}(\mathbf{x}) = \begin{bmatrix} 0 & -1 & y \\ 1 & 0 & -x \end{bmatrix} \quad \text{with} \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{3.48}$$

For infinite points $\mathbf{v}_\infty$, and lines $\boldsymbol{\lambda}$ in general, any component can be zero, and therefore one has to look for the largest element in $\mathbf{v}_\infty$ resp. $\boldsymbol{\lambda}$ and keep those two rows in $\mathsf{S}()$ that contain that element.

3.7 Conditioning of projective relations

The transition from $\mathbb{R}^2$ to $\mathbb{P}^2$ is given by relation (3.2) on page 12, which states that the projective representation $\mathbf{x}$ of a real 2D point $\mathbf{x}$ is simply obtained by adding a third coordinate with the value 1; i.e.

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \rightarrow \quad \mathbf{x} \sim \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} \boldsymbol{x}_O \\ x_H \end{pmatrix}.$$

Now suppose the point $\mathbf{x}$ is measured in a digital image with 2000×3000 pixels and is located far away from the origin of the coordinate frame. In this case the Euclidian part $\boldsymbol{x}_O$ of $\mathbf{x}$ will be in the order of 1000, whereas the homogenous part still is 1. This difference in order between the Euclidian and the homogenous part will cause enormous numerical problems if such projective points are used to derive other quantities; e.g. the projective transformation matrix H between two point sets.

These problems can be avoided easily if the projective entities are shifted and scaled prior to the computations. We will refer to this procedure as *conditioning*, as it improves the condition of the subsequent linear system of equations. This procedure is due to Hartley, who used this for computing the fundamental matrix; [Hartley 1995a]. He proposes to translate the set of points in the way that their centroid $\boldsymbol{x}_C$ is moved to the origin and then to scale the translated points

isotropically by $m = \sqrt{2}/s$, where s is the average distance of the points from $\boldsymbol{x}_C$. With $\mathbf{x}$ denoting the points in the original scale and $\bar{\mathbf{x}}$ denoting the conditioned points, this transformation can be represented as $\bar{\mathbf{x}} \sim \mathsf{H}\mathbf{x}$ with

$$\mathsf{H} = \begin{bmatrix} m & 0 & -mx_C \\ 0 & m & -my_C \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.49}$$

and due to equation (3.22) on page 22, $\mathsf{H}^{-\top}$ does the conditioning of 2D lines.

Note, however, that quantities which are obtained from these conditioned points do not refer to the original points. Therefore the conditioning matrix H has to be applied inversely on the results. For example, if the projective transformation matrix A between two point sets $\mathbf{x}$ and $\mathbf{y}$ is to be computed as $\mathbf{y} \sim \mathsf{A}\mathbf{x}$ and we apply conditioning matrices $\mathsf{H}_\mathbf{x}$ and $\mathsf{H}_\mathbf{y}$ respectively, then we actually solve the problem $\bar{\mathbf{y}} \sim \bar{\mathsf{A}}\bar{\mathbf{x}}$ with $\bar{\mathbf{x}} \sim \mathsf{H}_\mathbf{x}\mathbf{x}$ and $\bar{\mathbf{y}} \sim \mathsf{H}_\mathbf{y}\mathbf{y}$. The matrix A referring to the original points is obtained afterwards as $\mathsf{A} \sim {\mathsf{H}_\mathbf{y}}^{-1}\bar{\mathsf{A}}\mathsf{H}_\mathbf{x}$.

Chapter 4

A few Basics on Tensor Calculus

In this thesis only a very small portion of tensor calculus is needed, therefore we can keep this introduction short. For a more detailed discussion on this topic we refer to the text books on tensor calculus; e.g. [Spiegel 1959], [Klingbeil 1966] or [Joshi 1975].

The main difference between vector and tensor notation is that in the former a certain geometric entity, e.g. a point $\mathbf{x}$, is represented by a vector $\mathbf{x}$, whereas in the latter the point is directly represented by its components as x^i. Consequently, there is always some kind of index associated with the tensorial representation of an entity, and a **tensor** is simply an *indexed array of numbers*.

Indices and summation convention. There are two kinds of indices: sub-indices (called **co-variant**) and super-indices (called **contra-variant**). A tensor can have any number of co- and contra-variant indices. The number of these indices is called the **valence** of the tensor. A tensor of co-variant valence p and contra-variant valence q is termed a valence-(p,q) tensor and it is made up of n^{p+q} elements with n being the dimension of the underlying vector space; i.e. each index runs from 1 to n. In this context a scalar is a valency zero tensor, a vector is valence one tensor, a matrix is a valence two tensor and all tensors of higher valence are termed tensors in general.

From section 3.2 on page 12 we know that a vector of dimension 3 can either represent a point $\mathbf{x}$ or a line $\boldsymbol{\lambda}$ in $\mathbb{P}^2$. This duality is reflected in tensor notation in the following way: Points use a contra-variant index and lines use a co-variant index:

$$\mathbf{x} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = x^i \quad \text{and} \quad \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \lambda_i.$$

In vector notation a vector $\mathbf{x}$ sometimes needs to be transposed to operate correctly with other quantities. Such a transposition of a vector is not reflected in tensor notation, as this does not affect the components x^i. Therefore, for example the incidence of a point $\mathbf{x}$ with a line $\boldsymbol{\lambda}$, which is represented as $\mathbf{x}^\top \boldsymbol{\lambda} = 0$ in vector notation, must be realized in a different way in tensor notation. It is done using **Einstein's convention of summation**. This convention says that a sum is made up of all the same indices appearing as co- and contra-variant. So the incidence relation between a point $\mathbf{x}$ and a line $\boldsymbol{\lambda}$ is simply represented as

$$x^i \lambda_i := x^1\lambda_1 + x^2\lambda_2 + x^3\lambda_3 = 0.$$

If a matrix is to be represented in tensor notation it requires two indices, and there are no restrictions on the type of these indices: One can be co- and the other one can be contra-variant, or both can be of the same type. Which combination is the right one, depends on the property of the matrix. For example, a collineation H maps points to points. Keeping in mind Einstein's convention of summation, the map of a point $\mathbf{x}$ to a point $\mathbf{y}$, represented as $\mathbf{y} \sim \mathsf{H}\mathbf{x}$ in vector notation, is represented in tensor notation as

$$y^j \sim H^j_{\cdot i} x^i.$$

The *first index* (j) in $H^j_{\cdot i}$ refers to the rows of H and the *second index* (i) refers to the columns. The order of the indices is indicated by an additional $\cdot$ in $H^j_{\cdot i}$.

A correlation G maps a point $\mathbf{x}$ to a line $\boldsymbol{\lambda}$. Consequently the tensorial expression would be

$$\lambda_j \sim G_{ji} x^i.$$

Transposition of vectors is not reflected in tensor notation, as this is only needed in vector notation to make vectors operate correctly with other vectors or matrices. The transposition of matrices, however, must also be reflected in tensor notation, as the transposition of a matrix changes the property of the matrix. For example, if H maps points from system $\mathfrak{A}$ to points in system $\mathfrak{B}$, then $\mathsf{H}^\top$ maps lines from system $\mathfrak{B}$ to lines in system $\mathfrak{A}$. The transposition in tensor notation is realized by exchanging the respective indices; i.e.

$$\mathsf{H} = H^j_{\cdot i} \quad \rightarrow \quad \mathsf{H}^\top = H^{\cdot j}_i.$$

A tensorial expression, which only involves single or double indexed quantities can always be transferred to vector notation. Since the transposition of vectors is not reflected in tensor notation, a single indexed quantity x^i can either be interpreted as $\mathbf{x}$ or as $\mathbf{x}^\top$ during the transition to vector notation. So for example

$$s = G_{ij} x^i y^j,$$

can either be translated as

$$s = \mathbf{x}^\top \mathsf{G} \mathbf{y} \quad \text{or as} \quad s = \mathbf{y}^\top \mathsf{G}^\top \mathbf{x}.$$

Both transitions are of course valid.

As long as we are only dealing with single and double indexed quantities in tensor notation we actually need not apply tensor notation, because the same expressions can be represented in vector notation equally well and perhaps more accessible; and most of the relations in this thesis are of this kind. Tensor notation, however, together with Einstein's convention of summation has one strong benefit compared with vector notation. Tensor notation allows to deal with *operators, that work on more than two objects*.

A matrix is an operator of valence two; i.e. the input is either one or two vectors and the output is either a vector or a scalar. A tensor of valence three, is an operator on one, two or three vectors and the output is either a matrix, a vector or a scalar; i.e. the sum of the valencies of the input and the sum of the valencies of the output always gives the valence of the operator. The trifocal tensor, introduced in chapter 7 is such an operator of valence three, as it can handle points or lines from three different images at the same time. Therefore, certain relations can be expressed in a clear form only in tensor notation.

The fully antisymmetric valence-$(3,0)$ tensor ϵ_{ijk} For a general tensor of arbitrary valence, the terms symmetry and antisymmetry are defined only for pairs of similar indices. A tensor $A^{i\ k\ m}_{\cdot j \cdot l \cdot}$ is symmetric in the first and the third contra-variant index if $A^{i\ k\ m}_{\cdot j \cdot l \cdot} = A^{m\ k\ i}_{\cdot j \cdot l \cdot}$ or antisymmetric if $A^{i\ k\ m}_{\cdot j \cdot l \cdot} = -A^{m\ k\ i}_{\cdot j \cdot l \cdot}$. Symmetry and antisymmetry can only be defined for *similar* indices not for one co- and one contra-variant index, as it is shown in [Joshi 1975].

If symmetry or antisymmetry holds for all possible pairs of indices, then the respective tensor is said to be fully symmetric or fully antisymmetric. Of special interest is the **fully antisymmetric valence-$(3,0)$ tensor ϵ_{ijk}**. It is made up of $3^3 = 27$ elements, all of which are 0, except for the elements

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = -\epsilon_{132} = -\epsilon_{321} = -\epsilon_{213} = 1.$$

Using ϵ_{ijk} the cross-product of two vectors $\mathbf{x}$ and $\mathbf{y}$ can be expressed as

$$\mathbf{x} \times \mathbf{y} = \mathbf{z} = z_i = \epsilon_{ijk} x^j y^k. \tag{4.1}$$

Observe that the cross-product of two contra-variant coordinate vectors returns a co-variant coordinate vector. This might be surprising at first sight, but it is perfectly supported by section 3.5 on page 23 where we saw that the cross product is used for computing the line (a co-variant vector) joining two points (two contra-variant vectors). There we expressed the cross product using the so-called axiator, a 3×3 matrix $\mathsf{S}(\mathbf{x})$, as $\mathbf{z} = \mathsf{S}(\mathbf{x})\mathbf{y}$. From equation (4.1) it can be seen that the axiator $\mathsf{S}(\mathbf{x})$ can be expressed as

$$(\mathsf{S}(\mathbf{x}))_{ik} = \epsilon_{ijk} x^j. \tag{4.2}$$

If the cross-product of two co-variant vectors (corresponding to the intersection of two lines) is to be represented in tensor notation, we need the **fully antisymmetric valence-$(0,3)$ tensor ϵ^{ijk}**. And consequently the axiator $\mathsf{S}(\boldsymbol{\lambda})$ can be expressed as

$$(\mathsf{S}(\boldsymbol{\lambda}))^{ik} = \epsilon^{ijk} \lambda_j. \tag{4.3}$$

Chapter 5

Single Image Geometry

Prerequisite for any formulation of image geometry is the underlying camera model. The one, which is used throughout this thesis is the so-called *pinhole camera model* which will be described next.

5.1 The pinhole camera model

The purpose of the camera model is to describe the relation between the 2-dimensional image of an object taken with a camera and the 3D representation of the respective object. The central projection is a rather good model for this relation. A camera model that is based on the central perspective relation between object and image is called a **pinhole camera model**. In this model all the rays that come from the object pass through the same point in space, the so-called **projection center Z**. Afterwards the rays intersect the **image plane** Ψ, thereby generating the image. Note that in reality the projection center lies between object and image plane, hence the image is obtained mirrored and upside-down – the so-called *negative orientation*. Thus for easier handling the *positive orientation* is considered usually, with the image plane between **Z** and object, see figure 5.1 on the next page. The ray through the pinhole, i.e. **Z**, orthogonal to the image plane is called **principal axis**, its intersection point with the image plane is the **principal point** $\mathbf{p_0}$, the distance $\overline{\mathbf{Zp_0}}$ is called **principal distance** and will be denoted by c.

Physically the pinhole is realized as a system of lenses and the image plan is either a charge coupled device (CCD) sensor or a plane carrying a light sensitive film medium. Due to the inevitable limitations of these physical components any real image will *never* exactly correspond to a central perspective view but will have linear and non-linear distortions. The linear distortion can be incorporated easily in the model to be presented, the non-linear distortion, which is rather small for professional or photogrammetric cameras but may reach several pixels in a digital image of amateur cameras, is not part of this camera model.

As stated above in the first paragraph the camera model describes the relation between the image space and the object space. Both spaces are represented using proper coordinate systems. The image coordinate system will be described in more detail in the following. Actually we have to consider two kinds of image systems: the *camera coordinate system* and the *measurement system*.

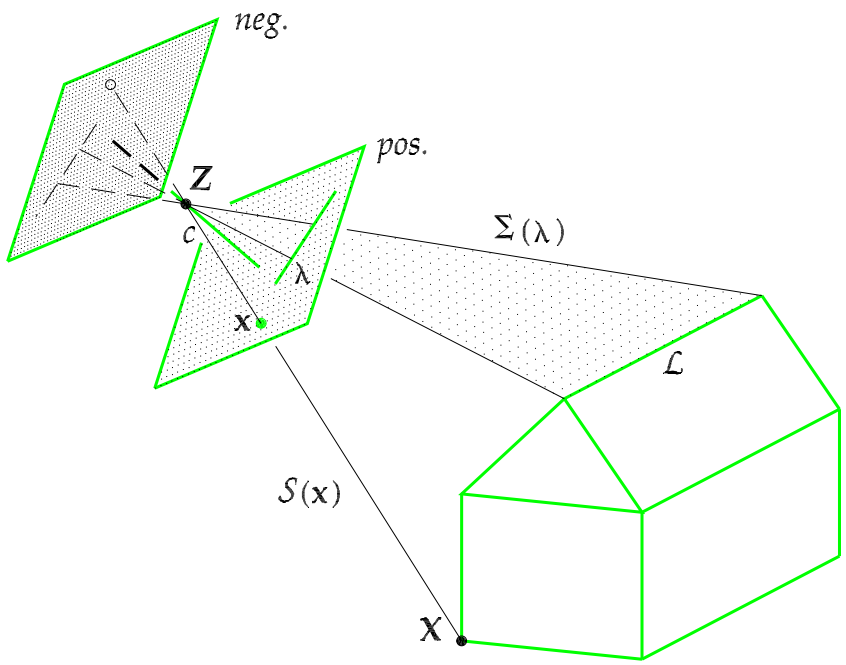

Figure 5.1: *Negative and positive orientation of an image; see section 5.1. Backprojection of a point and line in the image: projection ray $\boldsymbol{\mathcal{S}}(\mathbf{x})$ and projection plane $\boldsymbol{\Sigma}(\boldsymbol{\lambda})$; see section 5.4 on page 42.*

The **camera coordinate system** is defined as follows (see figure 5.2 on the facing page): The origin is incident with the principal point $\mathbf{p_0}$ and the first coordinate axis (x_c-axis) runs horizontally to the right. The second axis (y_c-axis) completes the system to be Cartesian and right-hand-oriented; i.e. y_c runs to the top. The **measurement system** is the coordinate system in which objects are actually observed. In the case of digital images it is the row/column system of the image matrix. In the case of analogous film images it is the system of the measuring device (analytical plotter or comparator). To link these two systems the measurements are translated[1] into the camera system, usually hereby affinity is also considered to cope with the linear distortion mentioned above.

In figure 5.2 the measurement system is drawn as an affine system with different scale k in y and a skew angle δ between the x- and y-axis. In this system the principal point has the coordinates (x_{p_0}, y_{p_0}). Following [Niini 1994] the transformation from the measurement system to the camera system can be expressed as

$$\begin{aligned} x_c &= (x - x_{p_0}) - k \sin\delta (y - y_{p_0}) \\ y_c &= k \cos\delta (y - y_{p_0}). \end{aligned} \tag{5.1}$$

Using $\alpha = k\cos\delta$ and $\beta = -k\sin\delta$ and homogenous coordinates this affine transformation

[1] If the measurements are only translated then the x_c-axis is aligned with the x-axis. Usually this is not a problem, since any rotation required in the image space can be made up by the rotation elements of the exterior orientation (see text), which will be *determined* afterwards. However, if these rotation elements of the exterior orientation are *observed* (e.g. using Inertial Measuring Units for aerial images; see e.g. [Cramer 2001]) then also a rotation from the measurement to the camera system can be necessary – especially when dealing with analog cameras. In such a case special marked points (so-called *fiducial marks*) are included in the image, whose camera coordinates are known from a laboratory calibration.

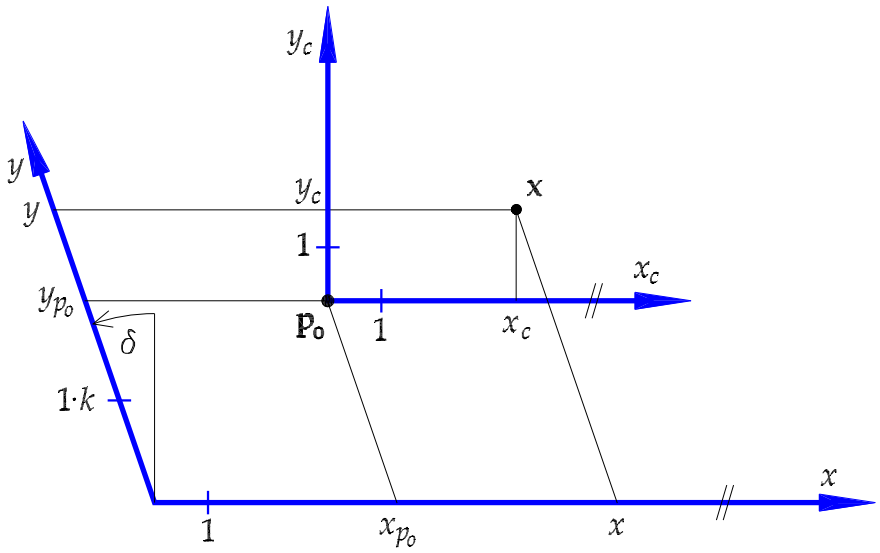

Figure 5.2: *The camera coordinate system and the affine measurement system.*

(5.1) can be represented by an upper-triangular matrix $\overline{\mathsf{C}}$

$$\overline{\mathsf{C}} = \begin{bmatrix} 1 & \beta & -x_{p_0} - \beta y_{p_0} \\ 0 & \alpha & -\alpha y_{p_0} \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.2}$$

as

$$\begin{pmatrix} x_c \\ 1 \end{pmatrix} = \overline{\mathsf{C}} \begin{pmatrix} x \\ 1 \end{pmatrix}. \tag{5.3}$$

5.2 The mapping of points – the projection matrix P

Now we can define a 3-dimensional image coordinate system $\{X_C, Y_C, Z_C\}$, which has the origin in the projection center $\mathbf{Z}$, the X_C- and Y_C-axis being parallel to the x_c- and y_c-axis and a Z_C-axis completing this right hand system; i.e. Z_C is orthogonal to the image plane pointing *away* from the object. This 3D image system is rotated by a 3-dimensional rotation matrix $\boldsymbol{R}$, which has determinant $+1$ due to the positive orientated systems involved, and translated by $\boldsymbol{Z}$ with respect to the object system. These 6 parameters (3 for $\boldsymbol{R}$ and 3 for $\boldsymbol{Z}$) are called the **exterior orientation** of the camera.

The image point $\boldsymbol{x}_c$ in the camera system for a given object point $\boldsymbol{X}$ can be found in the following way. The projection ray for this point is given by $\boldsymbol{X} - \boldsymbol{Z}$ in the object system and by $\boldsymbol{X}_C = \boldsymbol{R}^\top(\boldsymbol{X} - \boldsymbol{Z})$ in the 3D image coordinate system[2]. Following section 3.2 on page 12 $\boldsymbol{X}_C$ is the homogenous representation $\mathbf{x}_c$ of the image point $\boldsymbol{x}_c$, except that the homogenous part (i.e. the distance between the origin and the projective plane) refers to 1, whereas for the camera it needs to be the principal distance c. So in total the homogenous representation $\mathbf{x}_c$ of the image

[2]Other authors, e.g. [Hartley and Zisserman 2001] or [Heuel 2002], consider the rotation matrix without transposition; i.e. $\boldsymbol{X}_C = \boldsymbol{R}(\boldsymbol{X} - \boldsymbol{Z})$.

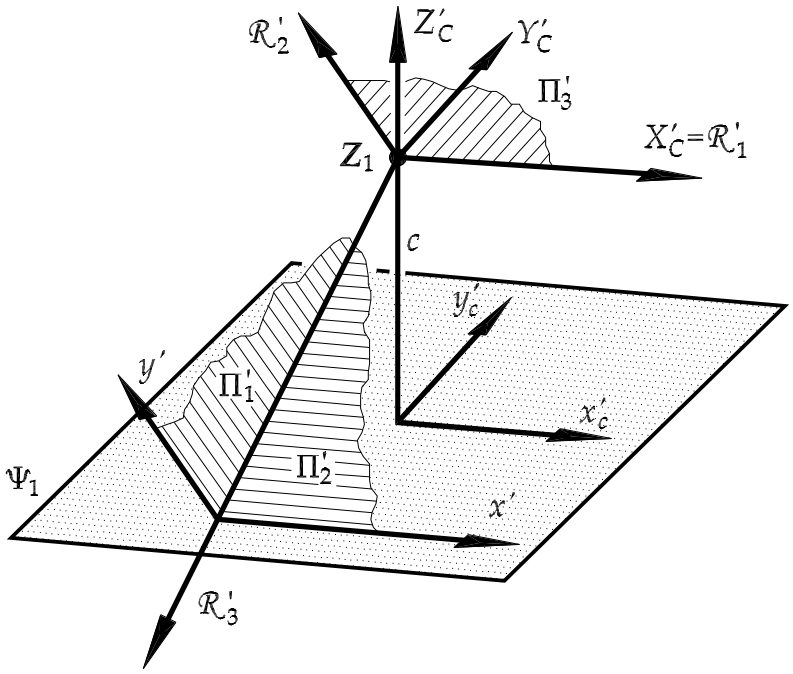

Figure 5.3: *The 3D image coordinate system $\{X_C, Y_C, Z_C\}$ of image Ψ with its principal planes $\mathbf{\Pi}_i$ and principal rays $\boldsymbol{\mathcal{R}}_i$ (see text).*

point in the camera system depending on the exterior orientation of the camera and the object point $\mathbf{X}$ is given by

$$\mathbf{x}_c \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{c} \end{bmatrix} R^\top (\boldsymbol{X} - \boldsymbol{Z}). \tag{5.4}$$

With the matrix $\overline{C}$ defined in equation (5.2) the central projection from the object point $\mathbf{X}$ to the image point $\mathbf{x}$ in the measurement system, both represented by homogenous vectors, can be expressed as

$$\mathbf{x} \sim \mathsf{P}\mathbf{X}, \tag{5.5}$$

with

$$\mathsf{P} \sim \mathsf{C}^{-1} R^\top [I_3, -\boldsymbol{Z}] \tag{5.6}$$

using

$$\mathsf{C} = \begin{bmatrix} 1 & \beta & -x_{p_0} - \beta y_{p_0} \\ 0 & \alpha & -\alpha y_{p_0} \\ 0 & 0 & -c \end{bmatrix}. \tag{5.7}$$

The matrix C will be referred to as **calibration matrix**, its inverse is given by

$$\mathsf{C}^{-1} = \frac{1}{c} \begin{bmatrix} c & -c\frac{\beta}{\alpha} & -x_{p_0} \\ 0 & \frac{c}{\alpha} & -y_{p_0} \\ 0 & 0 & -1 \end{bmatrix}. \tag{5.8}$$

The calibration matrix[3] C depends on the elements of the **interior orientation**. In Photogrammetry the interior orientation consists of two parts: (i) the coordinates (x_{p_0}, y_{p_0}) of the *principal point* $\mathrm{p}_0$and the *principal distance* c and (ii) *additional parameters* which model linear and non-linear image distortions. If these elements are known the images are called *calibrated* and *uncalibrated* otherwise. In the context of this thesis, the five elements $\{\alpha, \beta, x_{p_0}, y_{p_0}, c\}$ will be referred to as interior orientation, and non-linear image distortions are not considered.

In equation (5.6) the matrix P is introduced. It depends on the interior and exterior orientation of the camera and is termed **point projection marix**. P is a 3×4 matrix, thus represents a collineation from $\mathbb{P}^3$ to $\mathbb{P}^2$. P is required to have rank-3 otherwise not the entire image will be in the range of P. It is a homogenous matrix and thus has 11 degrees of freedom. Depending on given tasks it is helpful to consider different decompositions of P:

$$\mathsf{P} = \mathsf{M}\,[I_3, -\boldsymbol{Z}]\,, \tag{5.9}$$
$$\mathsf{P} = [\mathsf{M}, \mathbf{p}_4]\,, \tag{5.10}$$

with $\mathsf{M} = \mathsf{C}^{-1}R^\top$ being the homography from the plane at infinity $\boldsymbol{\Pi}_\infty$ to the image plane,

$$\mathsf{P} = [\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4]\,, \tag{5.11}$$
$$\mathsf{P} = \begin{bmatrix} \boldsymbol{\Pi}_1{}^\top \\ \boldsymbol{\Pi}_2{}^\top \\ \boldsymbol{\Pi}_3{}^\top \end{bmatrix}. \tag{5.12}$$

The last decomposition (5.12) is particularly interesting as the three $\boldsymbol{\Pi}_i$ represent 3D planes – the so-called **principal planes**; cf. [Papadopoulo and Faugeras 1998]. These are the three unique planes passing through $\boldsymbol{Z}$ shown in figure 5.3 on the facing page. We will see at the end of section 5.4, that these principal planes $\boldsymbol{\Pi}_i$ are the projection planes of the three canonical lines of $\mathbb{P}^2$. The plane $\boldsymbol{\Pi}_3$ parallel to the image plane is also called *focal plane*. The expression of $\boldsymbol{\Pi}_i$ in M and $\boldsymbol{Z}$ is simple

$$\boldsymbol{\Pi}_i \sim \begin{pmatrix} \boldsymbol{\Pi}_{i,H} \\ \Pi_{i,O} \end{pmatrix} \sim \begin{pmatrix} \mathsf{M}^\top \boldsymbol{i}_i \\ -\boldsymbol{Z}^\top \mathsf{M}^\top \boldsymbol{i}_i \end{pmatrix}. \tag{5.13}$$

Also of interest is the transformation of P if the object system is changed by a homography in $\mathbb{P}^3$ – represented by a regular homogenous 4×4 matrix H (i.e. $\overline{\mathbf{X}} \sim \mathsf{H}\mathbf{X}$). In order to keep the object-image-relation $\mathbf{x} \sim \mathsf{P}\mathbf{X} \sim \overline{\mathsf{P}}\overline{\mathbf{X}}$ the point projection matrix will then transform as

$$\overline{\mathsf{P}} \sim \mathsf{P}\mathsf{H}^{-1}. \tag{5.14}$$

Summarizing from [Hartley and Zisserman 2001] the projection matrix P has the properties shown in table 5.1 on the next page. From there we see that if the matrix M in $\mathsf{P} = [\mathsf{M}, \mathbf{p}_4]$ is singular the projection center lies infinitely far away. Since we deal with real perspective images this can not be the case, therefore we will assume the matrix M to be regular in general – except it is noted otherwise.

[3]Other authors refer to the inverse of C in relation (5.6) as the matrix K; e.g. [Hartley and Zisserman 2001] or [Heuel 2002].

Projection center The projection center $\mathbf{Z}$ is the 1-dimensional right null space of P; i.e. $\mathsf{P}\mathbf{Z} = \mathbf{0}$.

- M is regular (so-called *finite camera*): $\mathbf{Z} = \begin{pmatrix} -\mathsf{M}^{-1}\mathbf{p}_4 \\ 1 \end{pmatrix}$
- M is singular with rank 2 (so-called *camera at infinity*): $\mathbf{Z} = \begin{pmatrix} \mathbf{d} \\ 0 \end{pmatrix}$, with $\mathsf{M}\mathbf{d} = 0$.

Column vectors The column vectors $\mathbf{p}_i$ $(i = 1, 2, 3)$ are the *vanishing points* (maps of points at infinity) of the $X-$, $Y-$ and $Z-$axis of the object coordinate system; $\mathbf{p}_4$ is the image of the origin of the object coordinate system.

Row vectors The row vectors $\mathbf{\Pi}_i$ $(i = 1, 2, 3)$ represent the *principal planes*.

Principal point The principal point $\mathbf{p_0}$ can be computed as $\mathbf{p_0} \sim \mathsf{M}\mathsf{M}^\top \boldsymbol{i}_3$.

Principal axis The oriented vector $\boldsymbol{v}$ of the principal axis is given by $\boldsymbol{v} = -\det(\mathsf{M})\mathsf{M}^\top \boldsymbol{i}_3$. This vector is directed from the camera to the object; i.e. *opposite* to the Z_C-axis.

Table 5.1: *Properties of the point projection matrix* $\mathsf{P} = [\mathsf{M}, \mathbf{p}_4]$*; cf. [Hartley and Zisserman 2001].*

5.3 The mapping of lines – the projection matrix Q

Since the central projection represented by the pinhole camera model is a collineation, it is line preserving; cf. definition 3.2 on page 21. Consequently a relation for the mapping of lines, similar to the relation for points (5.5), is possible

$$\boldsymbol{\lambda} \sim \mathsf{Q}\boldsymbol{\mathcal{L}}. \tag{5.15}$$

Because 2D lines are represented as 3-dimensional homogenous vectors and 3D lines as 6-dimensional homogenous vectors (cf. sections 3.2 and 3.3.1 on page 12 and 17 respectively) the matrix Q is of dimension 3×6. Following [Förstner 2003] the matrix Q can be found in the following way.

The columns of Q are the image lines $\boldsymbol{\lambda}_i$ of the six canonical 3D lines $\boldsymbol{\mathcal{I}}_i$, $i = 1, \ldots, 6$, shown in table 3.3 on page 21, which are homogenously represented by the columns of I_6. Using the canonical points $\boldsymbol{I}_j$, $j = 1, \ldots, 4$, in $\mathbb{R}^4$ (see table 3.2 on page 17) these six lines can be represented as

$$\begin{aligned} \boldsymbol{\mathcal{I}}_1 &= \boldsymbol{I}_4 \vee \boldsymbol{I}_1 & \boldsymbol{\mathcal{I}}_4 &= \boldsymbol{I}_2 \vee \boldsymbol{I}_3 \\ \boldsymbol{\mathcal{I}}_2 &= \boldsymbol{I}_4 \vee \boldsymbol{I}_2 & \boldsymbol{\mathcal{I}}_5 &= \boldsymbol{I}_3 \vee \boldsymbol{I}_1 \\ \boldsymbol{\mathcal{I}}_3 &= \boldsymbol{I}_4 \vee \boldsymbol{I}_3 & \boldsymbol{\mathcal{I}}_6 &= \boldsymbol{I}_1 \vee \boldsymbol{I}_2. \end{aligned} \tag{5.16}$$

Consequently, the image lines $\boldsymbol{\lambda}_i$ of these six 3D lines must arise as the join of the maps of the respective canonical points $\boldsymbol{I}_j$. This leads to the form of the matrix Q given in the following proposition.

Proposition 5.1 (Line projection matrix.) *Let* $\mathsf{P} = [\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4]$ *be the point projection matrix. Then the line projection matrix* Q *is given by*

$$\mathsf{Q} \sim [\mathbf{p}_4 \times \mathbf{p}_1, \mathbf{p}_4 \times \mathbf{p}_2, \mathbf{p}_4 \times \mathbf{p}_1, \mathbf{p}_2 \times \mathbf{p}_3, \mathbf{p}_3 \times \mathbf{p}_1, \mathbf{p}_1 \times \mathbf{p}_2].$$

Especially if $\mathsf{P} \sim \mathsf{M}\,[I_3, -\boldsymbol{Z}]$ *with regular* M*, then* Q *turns into*

$$\mathsf{Q} \sim \mathsf{M}^{-\top}[-\mathsf{S}(\boldsymbol{Z}), I_3]\,.$$

Proof:
Let the line $\boldsymbol{\mathcal{L}}$ be defined by two points $\mathbf{A} = (A^1, A^2, A^3, A^4)^\top$ and $\mathbf{B} = (B^1, B^2, B^3, B^4)^\top$. Its map $\boldsymbol{\lambda}$ into the image is given by the join of the maps of $\mathbf{A}$ and $\mathbf{B}$. Using the 2D join relation (3.24) on page 23 this is represented as (observe the summation convention)

$$\boldsymbol{\lambda} \sim (\mathsf{P}\mathbf{A}) \times (\mathsf{P}\mathbf{B}) = A^i B^j \mathbf{p}_i \times \mathbf{p}_j.$$

Explicitly this sum reads as

$$\begin{aligned}(A^4B^1 - A^1B^4)\mathbf{p}_4 \times \mathbf{p}_1 + (A^4B^2 - A^2B^4)\mathbf{p}_4 \times \mathbf{p}_2 + (A^4B^3 - A^3B^4)\mathbf{p}_4 \times \mathbf{p}_3 \\ +(A^2B^3 - A^3B^2)\mathbf{p}_2 \times \mathbf{p}_3 + (A^3B^1 - A^1B^3)\mathbf{p}_3 \times \mathbf{p}_1 + (A^1B^2 - A^2B^1)\mathbf{p}_1 \times \mathbf{p}_2.\end{aligned}$$

The coefficients $(A^iB^j - A^jB^i)$ are exactly the coordinates of the Plücker vector $\boldsymbol{\mathcal{L}}$; cf. equation (3.14) on page 19. If M is regular, the cross products in the columns of Q can be simplified and combined using the axiator relation (B.12) on page 179.

□

A similar result for Q, but based on different Plücker line coordinates, can be found in [Carlsson 1997]. Another approach for deriving Q is suggested in [Faugeras and Luong 2001] and [Heuel 2002]. The row vectors of Q arise as the intersections of the three principal planes $\boldsymbol{\Pi}_1$, $\boldsymbol{\Pi}_2$, $\boldsymbol{\Pi}_3$ and thus are the dual 3D lines of the so-called **principal rays** $\boldsymbol{\mathcal{R}}_1$, $\boldsymbol{\mathcal{R}}_2$, $\boldsymbol{\mathcal{R}}_3$ depicted in figure 5.3 on page 38. Using relation (3.29) on page 24 for intersecting planes this is algebraically represented as

$$\mathsf{Q} \sim \begin{bmatrix} \boldsymbol{\Pi}_2 \wedge \boldsymbol{\Pi}_3 \\ \boldsymbol{\Pi}_3 \wedge \boldsymbol{\Pi}_1 \\ \boldsymbol{\Pi}_1 \wedge \boldsymbol{\Pi}_2 \end{bmatrix} \sim \begin{bmatrix} \boldsymbol{\mathcal{R}_1}^{*\top} \\ \boldsymbol{\mathcal{R}_2}^{*\top} \\ \boldsymbol{\mathcal{R}_3}^{*\top} \end{bmatrix}. \tag{5.17}$$

The principal rays can be expressed in the elements of the point projection matrix $\mathsf{P} \sim \mathsf{M}\,[I_3, -\boldsymbol{Z}]$ as

$$\boldsymbol{\mathcal{R}}_i = \begin{pmatrix} \boldsymbol{R}_{i,H} \\ \boldsymbol{R}_{i,O} \end{pmatrix} \sim \begin{pmatrix} \mathsf{M}^{-1}\boldsymbol{i}_i \\ \mathsf{S}(\boldsymbol{Z})\mathsf{M}^{-1}\boldsymbol{i}_i \end{pmatrix}. \tag{5.18}$$

Here the Plücker relation (3.15) on page 19 was used, where the projection center $\boldsymbol{Z}$ is a point on $\boldsymbol{\mathcal{R}}_i$ and $\mathsf{M}^{-1}\boldsymbol{i}_i$ is its direction. At the end of section 5.4 we will see that the principal rays are in fact the projection rays of the three canonical points of $\mathbb{P}^2$.

Finally we can ask for the transformation of Q if the object system is changed by a (point) homography in $\mathbb{P}^3$ – represented as a regular homogenous 4×4 matrix H (i.e. $\overline{\mathbf{X}} \sim \mathsf{H}\mathbf{X}$). The point homography defines a homography for 3D lines as $\overline{\boldsymbol{\mathcal{L}}} \sim \mathsf{H}_{\boldsymbol{\mathcal{L}}}\boldsymbol{\mathcal{L}}$; see proposition 3.1 on page 26. In order to keep the object-image-relation $\boldsymbol{\lambda} \sim \mathsf{Q}\boldsymbol{\mathcal{L}} \sim \overline{\mathsf{Q}}\overline{\boldsymbol{\mathcal{L}}}$ the line projection matrix will then transform as

$$\overline{\mathsf{Q}} \sim \mathsf{Q}{\mathsf{H}_{\boldsymbol{\mathcal{L}}}}^{-1}. \tag{5.19}$$

5.4 The back projection of points and lines

One of the main task of Photogrammetry and Computer Vision is to derive spatial information using given images. For this purpose at least two images from different view points of the same object are required. If only one image is given and no further information about the object is known the object can not be reconstructed uniquely. It is, nevertheless, important to know the set of 3D points resp. 3D lines that correspond to a particular given point $\mathbf{x}$ resp. line $\boldsymbol{\lambda}$ in the image. In the first case the set of 3D points is given by the projection ray $\boldsymbol{\mathcal{S}}(\mathbf{x})$ and in the second case the set of 3D lines is given by the projection plane $\boldsymbol{\Omega}(\boldsymbol{\lambda})$; see also figure 5.1 on page 36. Using the point and line projection matrices P and Q introduced in the previous sections, this so-called *back projection* of a point and a line can be expressed easily.

As it is shown in [Heuel 2002] the expression for the back projection of lines can be found in the following way. By considering the central projection (5.5) on page 38 as a collineation from $\mathbb{P}^3$ to $\mathbb{P}^2$ and by applying the relation (3.21) on page 22 for the dual collineation, we get

$$\mathsf{P}^\top \mathbf{x}^* \sim \mathbf{X}^*.$$

Since the dual of a 2D point $\mathbf{x}^*$ is a 2D line $\boldsymbol{\lambda}$ and the dual of 3D point $\mathbf{X}^*$ is a plane $\boldsymbol{\Sigma}$ we found the **back projection of a line $\boldsymbol{\lambda}$** to be

$$\boxed{\boldsymbol{\Sigma}(\boldsymbol{\lambda}) \sim \mathsf{P}^\top \boldsymbol{\lambda}.} \tag{5.20}$$

Note that with $\mathsf{P} \sim \mathsf{M}\,[I_3, -\boldsymbol{Z}]$ it is easy to see that the *normal vector of the projection plane* is given by

$$\boldsymbol{\Sigma}_H(\boldsymbol{\lambda}) \sim \mathsf{M}^\top \boldsymbol{\lambda} \tag{5.21}$$

or by

$$\boldsymbol{\Sigma}_H(\boldsymbol{\lambda}) \sim R\mathsf{C}^{-\top} \boldsymbol{\lambda} \tag{5.22}$$

in the elements of the interior and exterior orientation. Analogously by considering the dual of relation (5.15) we get

$$\mathsf{Q}^\top \boldsymbol{\lambda}^* \sim \boldsymbol{\mathcal{S}}^*.$$

Since the dual of a 2D line $\boldsymbol{\lambda}^*$ is a 2D point $\mathbf{x}$ and the dual of 3D line $\boldsymbol{\mathcal{S}}^*$ can be expressed as $\mathsf{D}\boldsymbol{\mathcal{S}}$ (using (3.18) on page 20), we found the **back projection of a point $\mathbf{x}$** to be

$$\boxed{\boldsymbol{\mathcal{S}}(\mathbf{x}) \sim \mathsf{Q}^{*\top} \mathbf{x},} \tag{5.23}$$

with

$$\mathsf{Q}^* = \mathsf{Q}\mathsf{D}. \tag{5.24}$$

And with $\mathsf{P} \sim \mathsf{M}\,[I_3, -\boldsymbol{Z}]$ it is easy to see that the *direction vector of the projection ray* is given by

$$\boldsymbol{S}_H(\mathbf{x}) \sim \mathsf{M}^{-1}\mathbf{x} \tag{5.25}$$

or by

$$S_H(\mathbf{x}) \sim R\mathsf{C}\mathbf{x} \tag{5.26}$$

in the elements of the interior and exterior orientation.

With these relations for the back projection it is easy to see, that the principal planes (the rows of P) are the projection planes of the canonical lines of the image, and that the principal rays (the rows of Q^*) are the projection rays of the canonical points of the image. For the interpretation of the canonical points and lines in $\mathbb{P}^2$ see table 3.1 on page 15.

Chapter 6

Two-View Geometry

In this chapter we consider the geometry of two images. Especially we will deal with the so-called *epipolar geometry*, which is enclosed in a 3×3 matrix, the *fundamental matrix*. At the end of this chapter we will look at relations between two images, which are induced by a plane and a line in 3D space. All these topics depend only on the so-called *relative orientation* of the images, which is described in the next section.

To distinguish the quantities associated with each image we will use the image number in our notation from now on. Partly we will do this by using the image number as a sub-script and partly by using the respective number of primes ($'$). So, for example $\mathbf{Z}_1$ is the projection center of image Ψ_1, and $\mathbf{x}''$ is a point in image Ψ_2.

6.1 The relative orientation of two images

At first we will define the term *relative orientation*.

Definition 6.1 (Relative orientation.) *A set of images is said to be relatively orientated if the projection rays of corresponding image points resp. the projection planes of corresponding image lines intersect in space. The term **relative orientation** is used since this condition is independent on the so-called **absolute orientation** of the set of images. In case of calibrated images the absolute orientation is a spatial similarity transformation, and in the case of uncalibrated images it is a 3D projective transformation.*

The relative orientation is of particular interest as it can be determined by the image observations alone – no information from the object space is required. Suppose the relative orientation of the images is given, then the object can be reconstructed up to a spatial similarity transformation if the interior orientation of the images is known, and up to a 3D projective transformation if the interior orientation of the images is unknown; see e.g. [Hartley and Zisserman 2001], where also the special case of an affine camera is discussed, for which reconstruction is possible up to a 3D affine transformation.

Proposition 6.1 (Degrees of freedom – calibrated images.) *The relative orientation of n calibrated images has 6n-7 degrees of freedom.*

Proof:
Each calibrated image has 6 degrees of freedom, which are the elements of the exterior orientation and 7 parameters are occupied by the absolute orientation (3 translations, 3 rotations and 1 scale).

□

Proposition 6.2 (Degrees of freedom – uncalibrated images.) *The relative orientation of n uncalibrated images has 11n-15 degrees of freedom.*

Proof:
Each uncalibrated image has 11 degrees of freedom, which are the elements of the exterior and interior orientation (i.e. the point projection matrix) and 15 parameters are occupied by the absolute projective orientation, which is represented by a homogenous 4×4 matrix.

□

Since two images taken from different view-points is the minimum requirement to gain information about the 3D environment, the relative orientation of two images is the most important one. Figure 6.1 depicts the relative orientation of two images Ψ_1 and Ψ_2 with different projection centers $\mathbf{Z}_1$ and $\mathbf{Z}_2$.

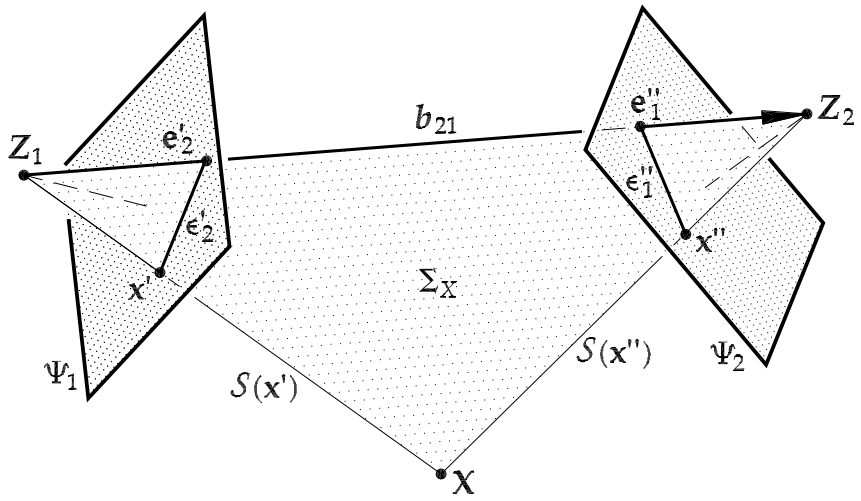

Figure 6.1: *The relative orientation of two images.*

Although the relative orientation is independent from the absolute orientation, the position and attitude of the images have to be expressed nevertheless with respect to some coordinate frame – which can be chosen arbitrarily. For the special case of the relative orientation of two calibrated images, which has 5 degrees of freedom, two ways for defining the coordinate frame have established:

- **The relative orientation of dependent images:** Here the projection center and the 3D image system of the first image defines the object frame. The second image is translated and rotated till all projection rays intersect in space. The length of the translation vector needs to be fixed in a proper way.

- **The relative orientation of independent images:** Here the two projection centers define one axis of the object frame. Both images are then rotated around their fixed centers till all projection rays intersect in space. One convenient rotation parameter of the first image is kept fix to yield 5 degrees of freedom.

Both definitions for the object frame can be used for the relative orientation of n images, in the way that, from the 3rd image on, all other images are positioned and rotated to yield intersecting projection rays. Usually we will use the relative orientation of dependent images, because the respective object frame is defined sole by one image – the *first* image – and we will consider the following definition.

Definition 6.2 (Relative orientation of dependent images.) *Whenever the projection centers $\boldsymbol{Z}_i$ and rotation matrices $\boldsymbol{R}_i$ of images Ψ_i (with $i = 1, 2, \ldots, n$) are referred to the relative orientation of dependent images (with respect to image Ψ_1) we will use a dag (†) to indicate this and it holds: $\boldsymbol{Z}_1^\dagger = \boldsymbol{0}$, $\boldsymbol{R}_1^\dagger = \boldsymbol{I}_3$ and $\boldsymbol{Z}_k^\dagger = \boldsymbol{R}_1^\top(\boldsymbol{Z}_k - \boldsymbol{Z}_1)$, $\boldsymbol{R}_k^\dagger = \boldsymbol{R}_1^\top \boldsymbol{R}_k$, with $k = 2, 3, \ldots, n$.*

6.1.1 Epipolar geometry

We look again at figure 6.1 on the preceding page. In each image there is one distinguished point – the so-called **epipole**, which is the map of the projection center of the other image; i.e. $\mathbf{e}_2'$ is the map of $\boldsymbol{Z}_2$ in image Ψ_1 and $\mathbf{e}_1''$ is the map of $\boldsymbol{Z}_1$ in image Ψ_2. Using the point projection matrix and decomposition (5.9) on page 39 an epipole $\mathbf{e}_k^{(j)}$ can generally be represented as

$$\mathbf{e}_k^{(j)} \sim \mathsf{P}_j \mathbf{Z}_k = \mathsf{M}_j(\boldsymbol{Z}_k - \boldsymbol{Z}_j). \tag{6.1}$$

The line joining the two projection centers is called **baseline**; b_{21} in figure 6.1. Any object point $\mathbf{X}$ together with this baseline defines a plane $\boldsymbol{\Sigma}_X$, called **epipolar plane**. The lines ε_2' and ε_1'' resulting from the intersection of $\boldsymbol{\Sigma}_X$ with Ψ_1 and Ψ_2 are called **epipolar lines**. Obviously all epipolar lines in one image must pass through the epipole in this image. Now we consider the point $\mathbf{x}'$, which is the map of the object point $\mathbf{X}$ in image Ψ_1. From figure 6.1 we see that the epipolar line ε_1'' can also be interpreted as the map of the projection ray $\boldsymbol{\mathcal{S}}(\mathbf{x}')$ into image Ψ_2 and that the corresponding point $\mathbf{x}''$ in image Ψ_2 has to lie on this epipolar line. This condition is called the **epipolar constraint**.

This epipolar constraint can be interpreted geometrically in different ways, see figure 6.1: (i) $\mathbf{x}'$ lies on ε_2' resp. $\mathbf{x}''$ lies on ε_1'', (ii) the projection rays $\boldsymbol{\mathcal{S}}(\mathbf{x}')$ and $\boldsymbol{\mathcal{S}}(\mathbf{x}'')$ intersect in space, or (iii) the baseline vector $\boldsymbol{Z}_2 - \boldsymbol{Z}_1$ and the direction vectors $\boldsymbol{S}_H(\mathbf{x}')$ and $\boldsymbol{S}_H(\mathbf{x}'')$ are coplanar. Of course, all interpretations yield the same algebraic representation for the epipolar constraint, which is given next.

6.1.2 The fundamental matrix

Algebraically the epipolar constraint is represented by a bilinear form according to the following theorem.

Theorem 6.1 (Fundamental matrix.) *Let two images Ψ_1 and Ψ_2 be given by their point projection matrices as $\mathsf{P}_i \sim \mathsf{M}_i\,[I_3, -\boldsymbol{Z}_i]$ with regular M_i and $i = 1, 2$. Further let $\{\mathbf{x}', \mathbf{x}''\}$ be a pair of corresponding image points. Then the* ***epipolar constraint*** *is given by*

$$\mathbf{x}'^{\top}\mathsf{F}_{12}\mathbf{x}'' = 0. \tag{6.2}$$

The matrix F_{12} is termed ***fundamental matrix*** *and is given by*

$$\mathsf{F}_{12} \sim \mathsf{M}_1{}^{-\top}\mathsf{S}(\mathbf{Z}_2 - \mathbf{Z}_1)\mathsf{M}_2^{-1}. \tag{6.3}$$

The relation between the interior and exterior orientation of the two images and the fundamental matrix is given by

$$\mathsf{F}_{12} \sim \mathsf{C}_1{}^{\top}R_1{}^{\top}\mathsf{S}(\mathbf{Z}_2 - \mathbf{Z}_1)R_2\mathsf{C}_2. \tag{6.4}$$

Proof:
Since we have different geometrical ways to proof the epipolar constraint, we choose the condition, that the projection rays $\boldsymbol{\mathcal{S}}(\mathbf{x}')$ and $\boldsymbol{\mathcal{S}}(\mathbf{x}'')$ of corresponding points in image Ψ_1 and image Ψ_2 intersect in space. These projection rays can be expressed using the line projection matrix Q of proposition 5.1 on page 40 and relation (5.23) on page 42 for the back projection of image points as $\boldsymbol{\mathcal{S}}(\mathbf{x}') \sim \mathsf{Q}_1^{*\top}\mathbf{x}'$ and $\boldsymbol{\mathcal{S}}(\mathbf{x}'') \sim \mathsf{Q}_2^{*\top}\mathbf{x}''$. The condition for the projection rays to intersect can be expressed by the coplanarity condition (equation (3.16) on page 20) of the representing Plücker vectors as $\boldsymbol{\mathcal{S}}(\mathbf{x}')^{\top}\boldsymbol{\mathcal{S}}^*(\mathbf{x}'') = 0 = \mathbf{x}'^{\top}\mathsf{Q}_1^*\mathsf{Q}_2{}^{\top}\mathbf{x}''$. This readily gives the epipolar constraint and the form (6.3) of the fundamental matrix F_{12}. The relation to the interior and exterior orientation is obtained using the equation (5.6) on page 38.

In [Niini 2000] a different derivation for the fundamental matrix is presented, which is also based on the intersection of the projection rays. Because it is an interesting approach it will be outlined next. The direction vector $\boldsymbol{S}_H(\mathbf{x}')$ of the projection ray is given by $\mathsf{M}_1^{-1}\mathbf{x}'$; see equation (5.23) on page 42. The condition for the object point $\mathbf{X}$ to lie on $\boldsymbol{S}(\mathbf{x}')$ can be expressed as $\boldsymbol{X} = \boldsymbol{Z}_1 + \mu_1\mathsf{M}_1^{-1}\mathbf{x}'$. For image Ψ_2 an analog expression holds and by equating these two expressions we get $\mu_1\mathsf{M}_1^{-1}\mathbf{x}' = \boldsymbol{Z}_2 - \boldsymbol{Z}_1 + \mu_2\mathsf{M}_2^{-1}\mathbf{x}''$, from which readily follows the epipolar constraint and the form of the fundamental matrix by multiplying with $\mathsf{S}(\boldsymbol{Z}_2 - \boldsymbol{Z}_1)$ and $\mathbf{x}'^{\top}\mathsf{M}_1{}^{-\top}$. This derivation is particularly interesting because it shows, that the relation between the two images turns into a homography (induced by the plane at infinity, see section 6.2.1 on page 53), if the two projection centers coincide; i.e. with $\boldsymbol{Z}_2 - \boldsymbol{Z}_1 = \mathbf{0}$ we get $\mathbf{x}' \sim \mathsf{M}_1\mathsf{M}_2^{-1}\mathbf{x}'' \sim \mathsf{H}_\infty\mathbf{x}''$.

□

Using definition 6.2 for the relative orientation of dependent images, relation (6.4) turns into

$$\mathsf{F}_{12} \sim \mathsf{C}_1{}^{\top}\mathsf{S}(\mathbf{Z}_2^{\dagger})R_2^{\dagger}\mathsf{C}_2. \tag{6.5}$$

For theorem 6.1 we assumed the projection matrices P_i to have regular M_i. If this requirement is not fulfilled, i.e. the projection centers lie at infinity, then F_{12} can be computed using formula (C.4) on page 184.

The fundamental matrix introduced in theorem 6.1 has the properties summarized in the following proposition; cf. [Hartley and Zisserman 2001].

Proposition 6.3 (Properties of the fundamental matrix.) *If the projection centers of the two images do not coincide, then* F_{12} *is a* rank-2 *homogenous matrix with 7 degrees of freedom. The epipoles* $\mathbf{e}'_2$ *and* $\mathbf{e}''_1$ *are the left and right nullspaces of* F_{12}*; i.e.*

$$\mathsf{F}_{12}{}^{\top}\mathbf{e}'_2 = 0, \quad \mathsf{F}_{12}\mathbf{e}''_1 = 0. \tag{6.6}$$

The epipolar lines $\boldsymbol{\varepsilon}'_2$ *and* $\boldsymbol{\varepsilon}''_1$ *for a pair of corresponding points* $\{\mathbf{x}', \mathbf{x}''\}$ *are obtained as*

$$\boldsymbol{\varepsilon}'_2 = \mathsf{F}_{12}\mathbf{x}'', \quad \boldsymbol{\varepsilon}''_1 = \mathsf{F}_{12}{}^{\top}\mathbf{x}'. \tag{6.7}$$

If the images are interchanged the corresponding fundamental matrix is

$$\mathsf{F}_{21} = \mathsf{F}_{12}{}^{\top}. \tag{6.8}$$

From relation (6.7) we see that F_{12} maps a point $\mathbf{x}''$ from image Ψ_2 to its corresponding epipolar line in image Ψ_1. This epipolar line represents the map of the projection ray of the point $\mathbf{x}''$. Therefore, the column vectors of F_{12} can be interpreted as the epipolar lines corresponding to the canonical points of image Ψ_2 or the maps of the projection rays of those canonical points. From section 5.4 on page 42 we know that these projection rays are the principal rays of image Ψ_2. Therefore we can give the following proposition.

Proposition 6.4 (The maps of the principal rays.) *The* i*th column vector of* F_{12} *represents the map of the* i*th principal ray* $\boldsymbol{\mathcal{R}}''_i$ *of image* Ψ_2 *into image* Ψ_1*. This line will be referred to as* $\boldsymbol{\rho}'_{2i}$*.*

Conversely, the i*th row vector of* F_{12} *represents the map of the* i*th principal ray* $\boldsymbol{\mathcal{R}}'_i$ *of image* Ψ_1 *into image* Ψ_2*, which will be referred to as* $\boldsymbol{\rho}''_{1i}$*.*

Note that the fundamental matrix is only of interest if the two projection centers do not coincide because otherwise $\mathsf{F}_{12} = \mathsf{0}$. In the proof of theorem 6.1 on the preceding page we saw that if the projection centers of the two images coincide, then the point sets in the two images are related by a homography induced by the plane at infinity (see section 6.2.1 on page 53).

Since the fundamental matrix maps points from the first image to (epipolar) lines in the second image, it represents a *projective correlation*. And since all epipolar lines are concurrent with the epipole $\mathbf{e}''_1$, it is a *singular* correlation, thus $|\mathsf{F}_{12}| = 0$. Since the fundamental matrix is made up of 9 elements but has only 7 degrees of freedom, F_{12} has to fulfil 2 constraints. One is the singularity constraint $|\mathsf{F}_{12}| = 0$ and the other one is the fixing of the scale since F_{12} is a homogenous matrix.

Computation of the fundamental matrix. The fundamental matrix fully represents the epipolar geometry of two uncalibrated images and therefore depends only on the relative orientation of the two images. Consequently F_{12} and the epipolar constraint, which is linear in F_{12}, provide a suitable means to determine linearly the relative orientation of two images if at least 8 corresponding pairs of image points are available – the so-called *8-point-algorithm*. For $n \geq 8$ corresponding pairs of points in two images this algorithm leads to a homogenous system of equations of the form[1]

$$\mathsf{A}\boldsymbol{f} = 0, \tag{6.9}$$

[1] Actually, A and $\boldsymbol{f}$ depend on projective entities, and therefore should be represented as A and $\mathbf{f}$. In the context of estimating projective entities based on an adjustment, where we do not apply any projective geometry at all, it is, however, preferable to consider the involved entities rather belonging to $\mathbb{R}^{n+1}$ than to $\mathbb{P}^n$. This way both worlds – projective geometry and parameter estimation, which involves adjustment theory and statistics – are clearly kept apart.

with an $n \times 9$ matrix A and $f = \mathrm{vec}(\mathsf{F}_{12})$. The ith row of A holds the derivatives of the epipolar constraint (6.2) for the ith point correspondence $\{\mathbf{x}'_i, \mathbf{x}''_i\}$ with respect to f and using the Kronecker-vec-relation (B.27) on page 180 the epipolar constraint can be written as

$$\mathbf{x}'^{\top} \mathsf{F}_{12} \mathbf{x}'' = 0 \quad \rightarrow \quad (\mathbf{x}'' \otimes \mathbf{x}')^{\top} f = 0,$$

and so the ith row of A is represented as

$$i_i{}^{\top} A = (\mathbf{x}''_i \otimes \mathbf{x}'_i)^{\top}. \tag{6.10}$$

The system (6.9) is numerically best solved using the singular value decomposition (SVD) (see appendix B.2 on page 178), which returns the solution $\overline{\mathsf{F}}_{12}$. This way, however, algebraic error is minimized and the singularity constraint of the fundamental matrix is generally not satisfied. It can be applied afterwards by computing the singular value decomposition of $\overline{\mathsf{F}}_{12}$ and setting the smallest singular value to zero. In this case that rank-2 matrix $\overline{\overline{\mathsf{F}}}_{12}$ is obtained, which minimizes the Frobenius norm $||\overline{\overline{\mathsf{F}}}_{12} - \overline{\mathsf{F}}_{12}||_{Frob}$; see [Hartley, Gupta, and Chang 1992].

The minimum case of 7 pairs of corresponding points can also be dealt with the singular value decomposition of system (6.9). The 2-dimensional nullspace $\mathcal{N}(\mathsf{A})$ is spanned by the two vectors f_1 and f_2. The solution $\overline{f}$ for the fundamental matrix can then be found as that linear combination of f_1 and f_2 which satisfies the singularity constraint. In general this will be up to three solutions.

For more and detailed analysis concerning the computation of the fundamental matrix see e.g. [Luong and Faugeras 1996], [Faugeras and Luong 2001] or [Hartley and Zisserman 2001].

Conditioning. As it was already mentioned in section 3.7 on page 28, it is inevitable to do a conditioning if measurements are introduced as projective quantities in the computations. The importance of conditioning the measurements can be seen from the elements of the matrix A, which are the derivatives of the epipolar constraint of each point correspondence with respect to f. These elements are products of *homogenous* point coordinates, one from each image. The homogenous part x_H is 1 for all real points, whereas the coordinates of the Euclidian part x_O can have any value. Suppose the points are measured in digital images with 2000×3000 pixels, then the Euclidian part will be in the order of 1000. Consequently the elements of A will vary between 1^2 and 1000^2, and we will encounter numerical problems in the solution of (6.9).

With the conditioning matrices presented in section 3.7 we can easily solve this problem. However, in this case the computed fundamental matrix, represented by the vector $\bar{\mathbf{f}}$, refers to the system of the conditioned measurements. The transformation back to the system of the original measurements is done in the following way.

If the points in each image are conditioned by different matrices, i.e. $\overline{\mathbf{x}}' \sim \mathsf{H}'\mathbf{x}'$, and $\overline{\mathbf{x}}'' \sim \mathsf{H}''\mathbf{x}''$, where $\mathbf{x}$ represents the original measurement and $\overline{\mathbf{x}}$ the conditioned one. Then the original fundamental matrix F_{12} is computed from the conditioned one $\overline{\mathsf{F}}_{12}$ by

$$\mathsf{F}_{12} \sim \mathsf{H}'^{\top} \overline{\mathsf{F}}_{12} \mathsf{H}''.$$

This transformation can also be performed in the following way. Let the original fundamental matrix be represented by the vector $\mathbf{f}$ and the conditioned fundamental matrix be represented by the vector $\bar{\mathbf{f}}$. Then it holds:

$$\mathbf{f} \sim \mathsf{H}_F \bar{\mathbf{f}},$$

where

$$\mathsf{H}_F \sim (\mathsf{H}'' \otimes \mathsf{H}')^\top.$$

Although the latter version requires much more computation time than the first version, it is particularly useful, because it enables the transformation of the covariance matrix $\mathbf{\Sigma}_{\bar{f}\bar{f}}$ of the conditioned fundamental matrix $\bar{\mathbf{f}}$ to the covariance matrix $\mathbf{\Sigma}_{ff}$ of the original fundamental matrix $\mathbf{f}$ by

$$\mathbf{\Sigma}_{ff} = \mathsf{H}_F \mathbf{\Sigma}_{\bar{f}\bar{f}} \mathsf{H}_F^\top.$$

Critical configurations. When computing the fundamental matrix we also have to take care of critical configurations or surfaces, which do not allow a unique solution for F_{12}. A detailed analysis of such critical situations is given in [Maybank 1993]. A summary is also included in [Hartley and Zisserman 2001]. The general critical surface is the ruled quadric[2], being non-degenerate (a hyperboloid of one sheet) or degenerate (like two planes, cones or cylinders). If the two projection centers and all the object points lie on such a ruled quadric up to 3 solutions for the fundamental matrix are possible. If all object points lie in a common plane then a two-parameter family of homogenous matrices is the general solution for F_{12}.

6.1.3 The essential matrix

If the interior orientation of the images is known, the fundamental matrix turns into the **essential matrix**

$$\mathsf{E}_{12} \sim \mathsf{R}_1{}^\top \mathsf{S}(\mathbf{Z}_2 - \mathbf{Z}_1) \mathsf{R}_2, \tag{6.11}$$

therefore E_{12} and F_{12} are related by

$$\mathsf{E}_{12} \sim \mathsf{C}_1{}^{-\top} \mathsf{F}_{12} \mathsf{C}_2{}^{-1}. \tag{6.12}$$

Using definition 6.2 on page 45 for the relative orientation of dependent images relation (6.11) turns into

$$\mathsf{E}_{12} \sim \mathsf{S}(\mathbf{Z}_2^\dagger) \mathsf{R}_2^\dagger. \tag{6.13}$$

Actually the essential matrix is a special instance of a fundamental matrix and thus has to have the properties shown in proposition 6.3 on page 47. Especially E_{12} is also a rank-2 homogenous matrix, but has 5 degrees of freedom as we already saw at the beginning of section 6.1. It has the following additional properties proven in [Huang and Faugeras 1989].

Proposition 6.5 (Singular values of the essential matrix.) *Any matrix E_{12} is an essential matrix if and only if its singular values $\{\sigma_1, \sigma_2, \sigma_3\}$ satisfy the following condition*

$$\sigma_1 = \sigma_2 \quad \textit{and} \quad \sigma_3 = 0. \tag{6.14}$$

[2] A ruled quadric is a quadric that contains a straight line.

Since the essential matrix is made up of 9 elements but only has 5 degrees of freedom, E_{12} has to satisfy 4 constraints. Again one is the fixing of the scale since E_{12} is a homogenous matrix, thus 3 constraints remain. Due to proposition 6.5, however, already the two singular value constraints are sufficient for a matrix to become essential. This apparent contradiction is resolved by observing that the $\sigma_1 = \sigma_2$ constraint is equivalent to two independent constraints; see [Faugeras and Luong 2001]. The two singular value constraints can be used when computing the essential matrix by minimizing algebraic error. After computing $\overline{\mathsf{E}}_{12}$ with the eight point algorithm presented in the previous section, that valid essential matrix $\overline{\overline{\mathsf{E}}}_{12}$, which minimizes $||\overline{\overline{\mathsf{E}}}_{12} - \overline{\mathsf{E}}_{12}||_{Frob}$ can be found by computing the SVD of $\overline{\mathsf{E}}_{12}$ with $\overline{\mathsf{E}}_{12} = \boldsymbol{U}\mathsf{S}\boldsymbol{V}^\top$ and $\mathrm{diag}(\mathsf{S}) = (\overline{\sigma}_1, \overline{\sigma}_2, \overline{\sigma}_3)^\top$, where $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \overline{\sigma}_3$. Then $\overline{\overline{\mathsf{E}}}_{12}$ is obtained as $\overline{\overline{\mathsf{E}}}_{12} = \boldsymbol{U}\mathsf{T}\boldsymbol{V}^\top$ with $\mathrm{diag}(\mathsf{T}) = (t, t, 0)^\top$, where $t = (\overline{\sigma}_1 + \overline{\sigma}_2)/2$.

If a valid essential matrix is to be determined using the Gauß-Helmert model (see section D.5 on page 195), which allows the minimization of reprojection error, then the $\sigma_1 = \sigma_2$ constraint cannot be used as a single constraint, as in the moment of enditeration the differentials of this constraint vanish. This can be shown using MATLAB and the respective listing of the m-file is included in section E.1 on page 199. As a consequence other 4 independent constraints have to be used. K. Rinner already investigated these constraints, [Rinner 1942], [Rinner 1956], [Rinner 1963], and he proposed the following set of constraints in [Rinner 1963].

Proposition 6.6 (Minimal set of constraints for the essential matrix.) *If the matrix* $\mathsf{E}_{12} \sim [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ *is an essential matrix with* $||\mathsf{E}_{12}||_{\mathrm{Frob}} = \sqrt{2}$*, then* E_{12} *satisfies the following constraints*

$$\begin{aligned} |\mathbf{e}_1|^2 + |\mathbf{e}_2|^2 + |\mathbf{e}_3|^2 &= 2, \\ |\mathbf{e}_1|^2 + |\mathbf{e}_1^\star|^2 &= 1, \\ |\mathbf{e}_2|^2 + |\mathbf{e}_2^\star|^2 &= 1, \\ |\mathbf{e}_3|^2 + |\mathbf{e}_3^\star|^2 &= 1, \end{aligned}$$

with $\mathsf{E}_{12}^\star = [\mathbf{e}_1^\star, \mathbf{e}_2^\star, \mathbf{e}_3^\star] = [\mathbf{e}_2 \times \mathbf{e}_3, \mathbf{e}_3 \times \mathbf{e}_1, \mathbf{e}_1 \times \mathbf{e}_2]$ *being the co-factor matrix of* E_{12}*. With these constraints the length of the baseline* $\boldsymbol{Z}_2^\dagger = \boldsymbol{Z}_2 - \boldsymbol{Z}_1$ *is set to 1.*

For the minimum case of 5 pairs of corresponding points, it is shown in [Faugeras and Maybank 1990] that up to 10 solutions for E_{12} are possible. J. Philip [Philip 1996] made serval numerical test which indicate that practically up to 6 real solutions for E_{12} are possible. He also presented a linear (i.e. unique) solution for 6 pairs of corresponding points.

6.1.4 Retrieving the basis vector and the rotation matrix from the essential matrix

Assume we are given the essential matrix E_{12} for two images Ψ_1 and Ψ_2, e.g. by determining it with the 8-point-algorithm of section 6.1.2, possibly by considering the constraints of section 6.1.3; and the scale is chosen so that $||\mathsf{E}_{12}||_{\mathrm{Frob}} = \sqrt{2}$; i.e. $||\mathbf{Z}_2^\dagger|| = 1$. Then we can extract the basis vector $\mathbf{Z}_2^\dagger$ and the rotation matrix $\boldsymbol{R}_2^\dagger$ up to a four-fold ambiguity in a first stage. In a second stage a unique solution (up to scale) can be obtained by using a real point correspondence. Several methods for the first step have been presented over the years, we will summarize two of them.

Using the singular value decomposition. With $\mathsf{E}_{12} = USV^\top$ we have; e.g. [Hartley and Zisserman 2001]:

$$\mathbf{Z}_2^\dagger = \boldsymbol{u}_3 \quad \text{or} \quad \overline{\mathbf{Z}}_2^\dagger = -\boldsymbol{u}_3,$$

and

$$R_2^\dagger = UWV^\top \quad \text{or} \quad \overline{R}_2^\dagger = UW^\top V^\top,$$

where $\boldsymbol{u}_3$ is the third column of U. The matrix W is given by

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Note that the rotation matrices obtained by this method may have negative determinant and since we defined the rotation matrix of an image to have positive determinant in section 5.2 on page 37 a multiplication by -1 may by necessary.

Using the co-factor matrix $\mathsf{E}_{12}^\star$. This following procedure does not use the singular value decomposition. The basis vector $\mathbf{Z}_2^\dagger$ is found by the cross product of any two columns of E_{12}. Since in principle this cross product can $\mathbf{0}$, it is recommend to compute it for all three possible pairs of columns and to choose the result with maximum length. The second solution $\overline{\mathbf{Z}}_2^\dagger$ is obtained by changing the sign of the first one. For the first solution of the basis vector $\mathbf{Z}_2^\dagger$ we get two solutions for the rotation matrix $R_2^\dagger$, which following [Rinner 1963] can be expressed as

$$R_2^\dagger = \mathsf{E}_{12}^\star - \mathsf{S}(\mathbf{Z}_2^\dagger)\mathsf{E}_{12} \quad \text{and} \quad \overline{R}_2^\dagger = \mathsf{E}_{12}^\star + \mathsf{S}(\mathbf{Z}_2^\dagger)\mathsf{E}_{12}.$$

These two solutions are explained by the sign-ambiguity of E_{12} (which, however, does not affect $\mathrm{adj}(\mathsf{E}_{12})$). For the second solution of the basis vector $\overline{\mathbf{Z}}_2^\dagger$ we get also two solutions for the rotation matrix, which are, however, essentially the same as for $\mathbf{Z}_2^\dagger$.

So, whatever method we choose, we always end up with two solutions for the basis vector and two solutions for the rotation matrix[3], producing four possible solutions for the pair of basis vector and rotation matrix: $\{\mathbf{Z}_2^\dagger, R_2^\dagger\}$, $\{\overline{\mathbf{Z}}_2^\dagger, R_2^\dagger\}$, $\{\mathbf{Z}_2^\dagger, \overline{R}_2^\dagger\}$, $\{\overline{\mathbf{Z}}_2^\dagger, \overline{R}_2^\dagger\}$. The geometrical interpretation of these four solutions is shown in figure 6.2 on the following page. They just correspond to the four possible combinations of positive and negative orientations of the two images; cf. figure 5.1 on page 36. We can, however, obtain a unique solution for the basis vector and the rotation matrix, by requesting a positive orientation for both images.

Requiring a unique solution. Following [Niini 1994] this can be done in the following way. If $\mathbf{x}'$ and $\mathbf{x}''$ are known corresponding image points of some point in space, with $x'_H = x''_H = 1$, then it holds:

$$\mu_1 \mathsf{C}_1 \mathbf{x}' = \mathbf{Z}_2^\dagger + \mu_2 R_2^\dagger \mathsf{C}_2 \mathbf{x}''.$$

Multiplying with $\mathsf{S}(\mathbf{Z}_2^\dagger)$ and requesting the same sign for the scalars μ_1 and μ_2 gives

$$\mathrm{sign}(\mathsf{S}(\mathbf{Z}_2^\dagger)\mathsf{C}_1\mathbf{x}') = \mathrm{sign}(\mathsf{S}(\mathbf{Z}_2^\dagger)R_2^\dagger\mathsf{C}_2\mathbf{x}''). \tag{6.15}$$

[3]Note, however, if the constraints of the essential matrix are not satisfied, only the rotation matrix returned from the singular value decomposition will be a valid rotation matrix; i.e. $R^{-1} = R^\top$.

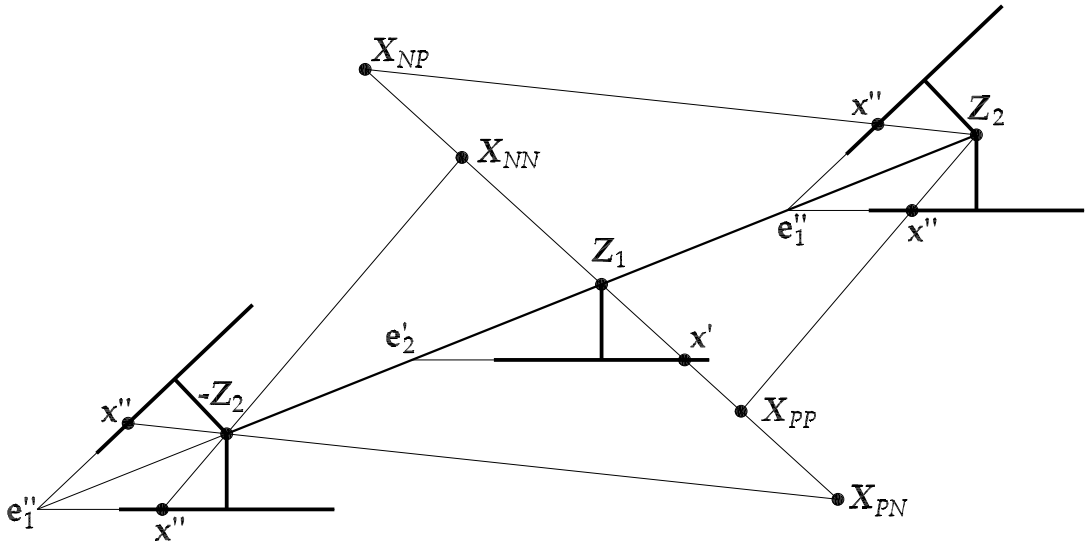

Figure 6.2: *The four solutions of the relative orientation of two images, which correspond to the four possible combinations of positive and negative orientations of the two images. The correct object point* $\mathbf{X}_{PP}$ *is only obtained if both images have positive orientation.*

This yields the correct rotation matrix and by requiring positive sign for the scalars μ_1 and μ_2, we finally get also the correct basis vector $\mathbf{Z}_2^{\dagger}$.

6.1.5 Historical notes

The representation of the relative orientation of two calibrated images using the bilinear form (6.2) on page 46 was rediscovered in photogrammetry several times; [von Sanden 1908], [Rinner 1942], [Thompson 1968] or [Stefanovic 1973] to name a few. The underlying matrix is given several names like *epipolar matrix*, *coplanarity matrix* or *correlation matrix* – if named at all. This matrix has 5 degrees of freedom and thus has to satisfy 4 constraints which were already presented in [Rinner 1942].

In computer vision the bilinear form of the relative orientation of calibrated images was introduced by [Longuet-Higgins 1981]. The underlying constraints and other properties of the *essential matrix* were investigated particularly by [Huang and Faugeras 1989] and [Maybank 1993]. The realization that a 3×3 matrix similar to the essential matrix can be used also for uncalibrated images, were made simultaneously by Faugeras [Faugeras 1992] and Hartley [Hartley 1992]. The term *fundamental matrix* was introduced by [Luong 1992].

6.2 Homographies and dual correlations between two images

In section 3.4 on page 20 we introduced the two basic projective transformations: *Homographies* (or collineations) and *correlations*. All transformations between projective spaces, that can be represented by matrices, belong to one of these two types; i.e.

mappings from points to points are *homographies*,
mappings from points to lines are *correlations*, and
mappings from lines to points are *dual correlations*.

We already encountered homographies in form of the point and line projection matrices P and Q in section 5, as well as a correlation in form of the fundamental matrix F_{12} in section 6.1.2. Also we will meet homographies and dual correlations in chapter 7, where we discuss the properties of the trifocal tensor. Therefore it is useful to summarize the properties of homographies and dual correlations separately.

6.2.1 Homographies between two images induced by a plane

We consider the mapping from a first image Ψ_1 to a second image Ψ_2 shown in figure 6.3. A point $\boldsymbol{x}'$ in Ψ_1 is back-projected into space, the respective projection ray $\boldsymbol{\mathcal{S}}(\boldsymbol{x}')$ is intersected with a plane $\boldsymbol{\Omega}$ and then the resulting intersection point $\mathbf{X}$ is projected into the second image Ψ_2 yielding the point $\boldsymbol{x}''$. Algebraically this mapping can be represented by a homogenous 3×3 matrix H_Ω and we will refer to such kind of mapping as a **homography from image Ψ_1 to image Ψ_2 induced by the plane $\boldsymbol{\Omega}$**.

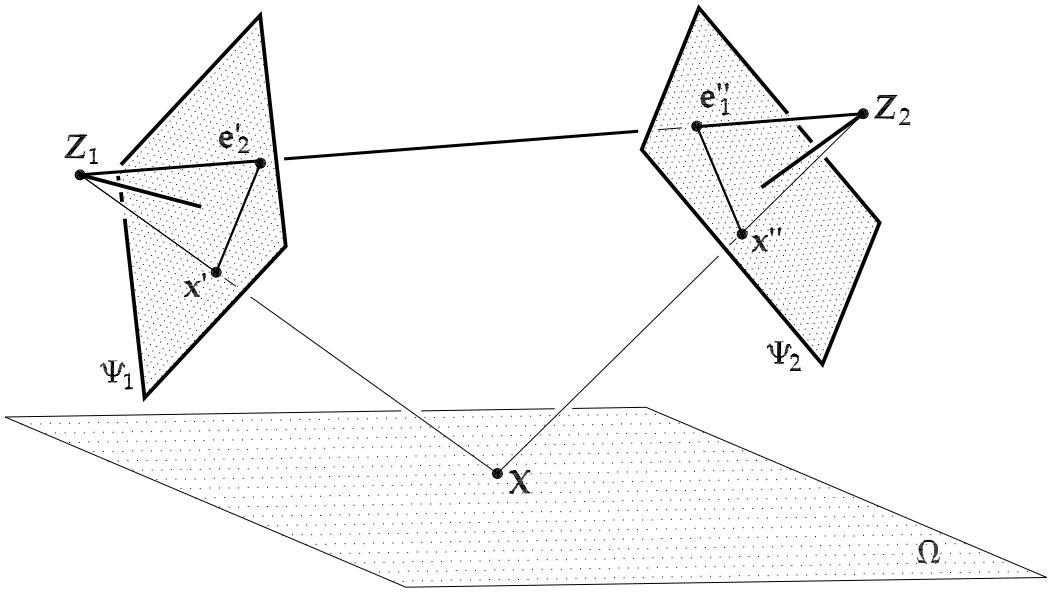

Figure 6.3: *The homography from image Ψ_1 to image Ψ_2 induced by the plane $\boldsymbol{\Omega}$.*

Proposition 6.7 (Homography induced by a plane.) *Let two images Ψ_1 and Ψ_2 be given by their point projection matrices $\mathsf{P}_1 \sim \mathsf{M}_1\,[I_3, -\boldsymbol{Z}_1]$ and $\mathsf{P}_2 \sim \mathsf{M}_2\,[I_3, -\boldsymbol{Z}_2]$ (with M_1 and M_2 regular), and let further a plane $\boldsymbol{\Omega}$ in 3D space be represented as $\boldsymbol{\Omega} \sim (\boldsymbol{\Omega}_H{}^\top, \Omega_O)^\top$. Then the homography H_Ω from image Ψ_1 to image Ψ_2 induced by the plane $\boldsymbol{\Omega}$ is given by*

$$\mathsf{H}_\Omega \sim \mathsf{M}_2(dI_3 + (\boldsymbol{Z}_2 - \boldsymbol{Z}_1)\boldsymbol{\Omega}_H{}^\top)\mathsf{M}_1^{-1} \quad \textit{with} \quad d = \boldsymbol{Z}_1{}^\top\boldsymbol{\Omega}_H + \Omega_O.$$

Proof:
Using the equations (5.23), (3.33) and (5.5) the projection ray $\boldsymbol{\mathcal{S}}(\mathbf{x}')$, its intersection point with $\boldsymbol{\Omega}$ and the projection $\mathbf{x}''$ into the second image can be represented algebraically as

$$\mathbf{x}'' \sim \mathsf{P}_2 \overline{\Pi}(\boldsymbol{\Omega}) \mathsf{Q}_1^{*\top} \mathbf{x}'.$$

By applying proposition 5.1 for the line projection matrix Q_1 and equation (B.10) for the arising cross products, the proposition follows immediately.

□

The matrix H_Ω of the homography induced be a general plane $\boldsymbol{\Omega}$ will be regular in general. For special cases, however, H_Ω will be singular. Depending on its rank the matrix H_Ω will have different properties.

Proposition 6.8 (Ranks of plane homographies.) *In general* $\mathrm{rank}(\mathsf{H}_\Omega) = 3$ *and the epipoles* $\mathbf{e}_2'$ *(in image* Ψ_1*) and* $\mathbf{e}_1''$ *(in image* Ψ_2*) are related by*

$$\mathbf{e}_1'' \sim \mathsf{H}_\Omega \mathbf{e}_2'. \tag{6.16}$$

Rank$(\mathsf{H}_\Omega) = 2$*, if the homography plane* $\boldsymbol{\Omega}$ *contains the projection center* $\mathbf{Z}_2$ *of image* Ψ_2*.*
Rank$(\mathsf{H}_\Omega) = 1$*, if the homography plane* $\boldsymbol{\Omega}$ *contains the projection center* $\mathbf{Z}_1$ *of image* Ψ_1*.*
Rank$(\mathsf{H}_\Omega) = 0$*, if the homography plane* $\boldsymbol{\Omega}$ *contains the coinciding projection centers* $\mathbf{Z}_1 = \mathbf{Z}_2$ *.*
Depending on its rank*,* H_Ω *has the ranges and nullspaces shown in table 6.1.*

$\mathrm{rank}(\mathsf{H}_\Omega)$	$\mathcal{R}(\mathsf{H}_\Omega)$	$\mathcal{N}(\mathsf{H}_\Omega)$
3	$\mathbb{P}^2$	$\{\}$
2	$\boldsymbol{\sigma}'' \sim \mathsf{M}_2^{-\top} \boldsymbol{\Omega}_H$	$\mathbf{e}_2'$
1	$\mathbf{e}_1''$	$\boldsymbol{\sigma}' \sim \mathsf{M}_1^{-\top} \boldsymbol{\Omega}_H$
0	$\{\}$	$\mathbb{P}^2$

Table 6.1: *Range* $\mathcal{R}(\mathsf{H}_\Omega)$ *and nullspace* $\mathcal{N}(\mathsf{H}_\Omega)$ *of the plane homography* H_Ω *depending on its rank; cf. the figures 6.4 and 6.5. The images* Ψ_1 *resp.* Ψ_2 *are given as* $\mathsf{P}_1 \sim \mathsf{M}_1 \left[I_3, -\mathbf{Z}_1\right]$ *resp.* $\mathsf{P}_2 \sim \mathsf{M}_2 \left[I_3, -\mathbf{Z}_2\right]$*;* $\mathbf{e}_2'$ *resp.* $\mathbf{e}_1''$ *are their epipoles. The plane* $\boldsymbol{\Omega}$ *is given as* $\boldsymbol{\Omega} \sim (\boldsymbol{\Omega}_H{}^\top, \Omega_O)^\top$ *and the lines* $\boldsymbol{\sigma}'$ *and* $\boldsymbol{\sigma}''$ *are the intersection lines of* $\boldsymbol{\Omega}$ *with* Ψ_1 *and* Ψ_2*.*

Proof:
First we proof the dependence of $\mathrm{rank}(\mathsf{H}_\Omega)$ on the incidence of the projection centers $\mathbf{Z}_1$ resp. $\mathbf{Z}_2$ with the plane $\boldsymbol{\Omega}$. For this we just have to consider the middle part of H_Ω because the enclosing M matrices are assumed to be regular. Setting $|\mathsf{H}_\Omega| = 0$ gives $|dI_3 + (\mathbf{Z}_2 - \mathbf{Z}_1)\boldsymbol{\Omega}_H{}^\top| = |\overline{\mathsf{H}}| = 0$. Observe that the left side of this equation can be interpreted as the characteristic polynomial of the dyad $-(\mathbf{Z}_2 - \mathbf{Z}_1)\boldsymbol{\Omega}_H{}^\top$ with eigen-value d. By expanding the determinant by one column, this equation reduces to

$$d^2(d + \boldsymbol{\Omega}_H{}^\top(\mathbf{Z}_2 - \mathbf{Z}_1)) = 0.$$

It goes without saying that $d_1 = d_2 = 0$ is a solution with multiplicity 2 and the resulting matrix $\overline{\mathsf{H}}_1 = \overline{\mathsf{H}}_2$ has rank-1 (since only the dyad remains). Since the solution $d_3 = -\boldsymbol{\Omega}_H{}^\top(\mathbf{Z}_2 -$

$\mathbf{Z}_1$) is unique, the resulting matrix $\overline{\mathsf{H}}_3$ has to have rank-2; see inequality (B.4) on page 177 which relates the multiplicity of an eigenvalue and the resulting rank defect. Then, with d from Proposition 6.7, the dependence of $\mathrm{rank}(\mathsf{H}_\Omega)$ on the incidence of the projection centers $\mathbf{Z}_1$ resp. $\mathbf{Z}_2$ with the plane $\boldsymbol{\Omega}$ follows directly. Thus, it only remains the rank-0 case. It is easy to see, that $\overline{\mathsf{H}} = \mathsf{0}_3$ if $d = 0$ and $\mathbf{Z}_1 = \mathbf{Z}_2$.

If $\mathrm{rank}(\mathsf{H}_\Omega) = 3$, then it follows immediately that $\mathbf{e}_1'' \sim \mathsf{H}_\Omega \mathbf{e}_2'$ because the epipoles $\mathbf{e}_2'$ resp. $\mathbf{e}_1''$ are defined by $\mathbf{e}_2' \sim \mathsf{M}_1(\boldsymbol{Z}_2 - \boldsymbol{Z}_1)$ resp. $\mathbf{e}_1'' \sim \mathsf{M}_2(\boldsymbol{Z}_1 - \boldsymbol{Z}_2)$.

Next we proof the range and nullspaces of H_Ω by utilizing the proven incidence relations of the projection centers with the plane $\boldsymbol{\Omega}$. Using the relation (B.10) on page 179 for the product of two axiators, a rank-2 homography matrix can be decomposed as $\mathsf{H}_\Omega \sim \mathsf{M}_2\mathsf{S}(\boldsymbol{\Omega}_H)\mathsf{S}(\boldsymbol{Z}_2 - \boldsymbol{Z}_1){\mathsf{M}_1}^{-1}$. With this decomposition it is easy to see that $\mathsf{H}_\Omega \mathbf{e}_2' = \mathbf{0}$ and this is the only possible nullspace as $\mathrm{rank}(\mathsf{H}_\Omega) = 2$. Using relation (B.12) on page 179 it follows that all mapped points $\mathbf{x}''$ have to lie on the line $\boldsymbol{\sigma}'' \sim {\mathsf{M}_2}^{-\top}\boldsymbol{\Omega}_H$. Similarly for $\mathrm{rank}(\mathsf{H}_\Omega) = 1$ the homography matrix can be decomposed as $\mathsf{H}_\Omega \sim \mathsf{M}_2(\boldsymbol{Z}_2 - \boldsymbol{Z}_1){\boldsymbol{\Omega}_H}^\top {\mathsf{M}_1}^{-\top}$, which can also be written as $\mathsf{H}_\Omega \sim \mathbf{e}_1''\boldsymbol{\sigma}'^\top$.

□

In the figures 6.4 and 6.5 rank-2 and rank-1 homographies are depicted; cf. table 6.1. Note that the inverse homography (i.e. the mapping of the points in image Ψ_2 to the points in image Ψ_1) can be derived from Proposition 6.7 on page 53 by exchanging the roles of the two images. So even for singular homographies the map in the opposite direction can be expressed.

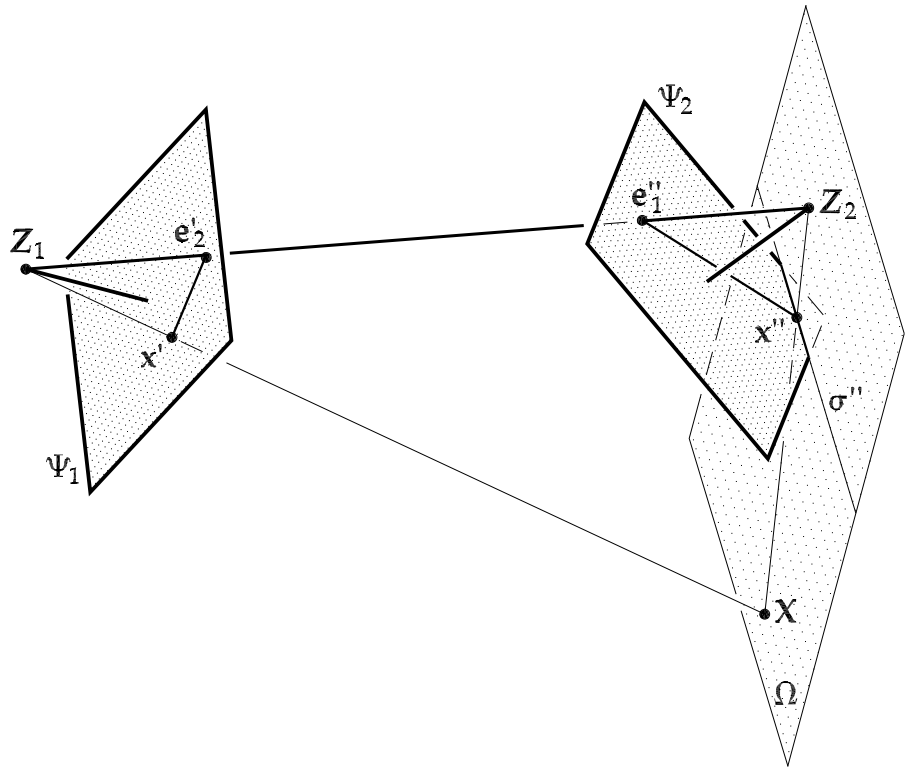

Figure 6.4: *A rank-2 homography from image Ψ_1 to image Ψ_2.*

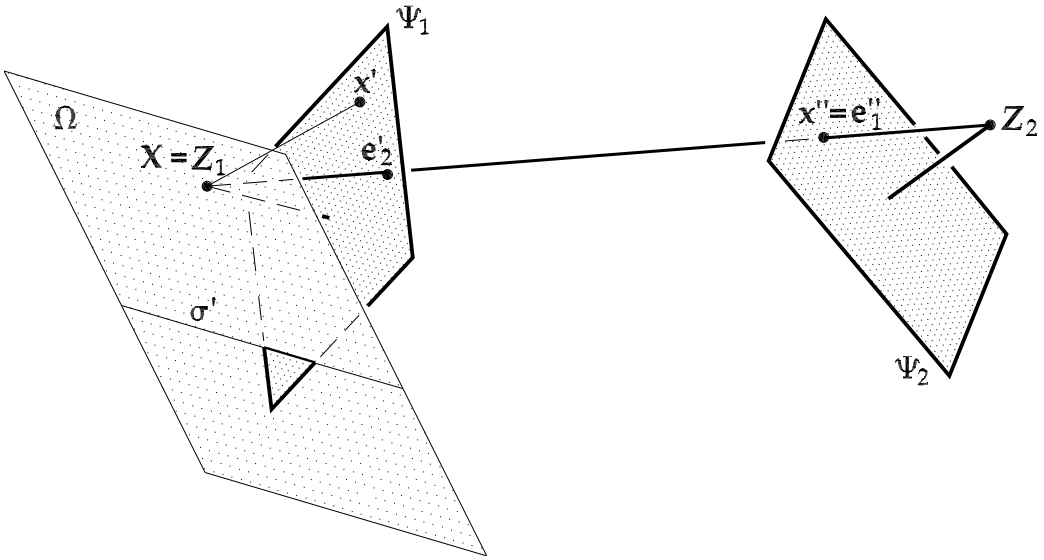

Figure 6.5: *A rank-1 homography from image Ψ_1 to image Ψ_2.*

A special homography is the one induced by the plane at infinity $\mathbf{\Pi}_\infty$ – denoted by H_∞. With $\mathbf{\Pi}_\infty = (0,0,0,1)^\top$ (see section 3.3) it is easy to see that this homography is represented by

$$\mathsf{H}_\infty \sim \mathsf{M}_2\mathsf{M}_1{}^{-1}. \tag{6.17}$$

The same homography H_∞ also results if the two projection centers $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ coincide. Therefore H_∞ can be interpreted in two ways: (i) H_∞ represents the homography induced by the plane at infinity, which turns out to be independent on the projection centers; equation (6.17). (ii) H_∞ represents the homography from image Ψ_1 to image Ψ_2 in case the projection centers coincide, which is then independent on the inducing plane $\mathbf{\Omega}$, provided the projection centers do not lie in $\mathbf{\Omega}$.

Finally, we can give the relation between a plane homography and the fundamental matrix of two images.

Proposition 6.9 (Plane homography and the fundamental matrix.) *Let H_Ω be the homography from image Ψ_1 to image Ψ_2 induced by a plane $\mathbf{\Omega}$, with* $\operatorname{rank}(\mathsf{H}_\Omega) > 1$. *Then the fundamental matrix F_{12} between these two images (with epipolar constraint* (6.2)*) is given by*

$$\mathsf{F}_{12} \sim \mathsf{H}_\Omega{}^\top \mathsf{S}(\mathbf{e}''_1).$$

Proof:
The epipolar line $\boldsymbol{\varepsilon}''_1$ in the second image Ψ_2 for a given point $\mathbf{x}'$ is given by $\boldsymbol{\varepsilon}''_1 \sim \mathsf{F}_{12}{}^\top \mathbf{x}'$; (see equation (6.7) on page 47).

A point $\mathbf{y}''$ on the epipolar line $\boldsymbol{\varepsilon}''_1$ can be found as $\mathbf{y}'' \sim \mathsf{H}_\Omega \mathbf{x}'$. By joining $\mathbf{y}''$ and $\mathbf{e}''_1$ we obtain the epipolar line $\boldsymbol{\varepsilon}''_1$. Using equation (3.24) on page 23 this is represented algebraically as $\boldsymbol{\varepsilon}''_1 \sim \mathsf{S}(\mathbf{e}''_1)\mathbf{y}'' \sim \mathsf{S}(\mathbf{e}''_1)\mathsf{H}_\Omega \mathbf{x}'$. The comparison of this expression for $\boldsymbol{\varepsilon}''_1$ with $\boldsymbol{\varepsilon}''_1 \sim \mathsf{F}_{12}{}^\top \mathbf{x}'$ gives the result. For the epipolar line $\boldsymbol{\varepsilon}''_1$ to be valid it is required that $\mathbf{y}'' \neq \mathbf{e}''_1$. This is true as long as $\mathbf{x}' \neq \mathbf{e}'_2$ and $\operatorname{rank}(\mathsf{H}_\Omega) > 1$; cf. table 6.1 on page 54.

□

6.2.2 Dual correlations between two images induced by a 3D line

Now we consider another mapping from a first image Ψ_1 to a second image Ψ_2, which is shown in figure 6.6. A line $\boldsymbol{\lambda}'$ in Ψ_1 is back-projected into space, the respective projection plane $\boldsymbol{\Sigma}(\boldsymbol{\lambda}')$ is intersected with a line $\boldsymbol{\mathcal{L}}$ and then the resulting intersection point $\mathbf{X}$ is projected into the second image Ψ_2 yielding the point $\mathbf{x}''$. Algebraically this mapping can be represented by a homogenous 3×3 matrix $\mathsf{G}_{\mathcal{L}}$ and we will refer to such kind of mapping as a **dual correlation from image Ψ_1 to image Ψ_2 induced by the 3D line $\boldsymbol{\mathcal{L}}$**.

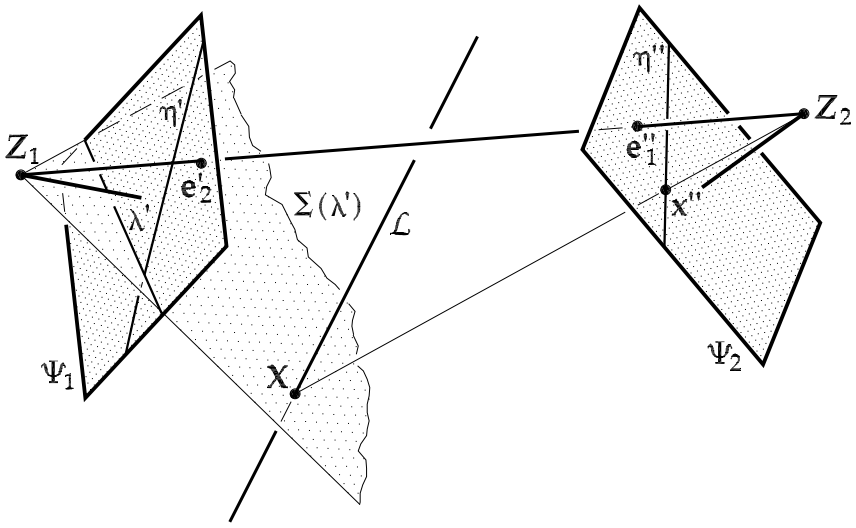

Figure 6.6: *The dual correlation from image Ψ_1 to image Ψ_2 induced by the line $\boldsymbol{\mathcal{L}}$.*

Proposition 6.10 (Dual correlation induced by a 3D line.) *Let two images Ψ_1 and Ψ_2 be given by their point projection matrices $\mathsf{P}_1 \sim \mathsf{M}_1 [I_3, -\boldsymbol{Z}_1]$ and $\mathsf{P}_2 \sim \mathsf{M}_2 [I_3, -\boldsymbol{Z}_2]$ (with M_1 and M_2 regular), and let further a line $\boldsymbol{\mathcal{L}}$ in 3D space be represented as $\boldsymbol{\mathcal{L}}^\top \sim (\boldsymbol{L}_H{}^\top, \boldsymbol{L}_O{}^\top)$. Then the dual correlation $\mathsf{G}_{\mathcal{L}}$ from image Ψ_1 to image Ψ_2 induced by the line $\boldsymbol{\mathcal{L}}$ is given by*

$$\mathsf{G}_{\mathcal{L}} \sim \mathsf{M}_2 \Big(\mathsf{S}(\boldsymbol{L}_O) + \boldsymbol{Z}_2 \boldsymbol{L}_H{}^\top - \boldsymbol{L}_H \boldsymbol{Z}_1{}^\top \Big) \mathsf{M}_1{}^\top.$$

Proof:
Using the equations (5.20), (3.33) and (5.5) the projection plane $\boldsymbol{\Sigma}(\boldsymbol{\lambda}')$, its intersection point with $\boldsymbol{\mathcal{L}}$ and the projection $\mathbf{x}''$ into the second image can be represented algebraically as

$$\mathbf{x}'' \sim \mathsf{P}_2 \boldsymbol{\Gamma}^\top(\boldsymbol{\mathcal{L}}) \mathsf{P}_1{}^\top \boldsymbol{\lambda}',$$

which directly gives the proposition.

□

Since only 3D points that lie on $\boldsymbol{\mathcal{L}}$ are projected into image Ψ_2, this kind of mapping will be of rank-2 in general. The following proposition goes more into detail.

Proposition 6.11 (Ranks of dual line correlations.) *In general $\operatorname{rank}(\mathsf{G}_{\mathcal{L}}) = 2$. The right ($\boldsymbol{\eta}'$) resp. left ($\boldsymbol{\eta}''$) nullspace of $\mathsf{G}_{\mathcal{L}}$ is the map of the line $\boldsymbol{\mathcal{L}}$ into image Ψ_1 resp. Ψ_2.*

Rank$(\mathsf{G}_{\mathcal{L}}) = 1$, *if the correlation line* $\boldsymbol{\mathcal{L}}$ *contains at least one of the projection centers of the two images, with* $\mathbf{Z}_1 \neq \mathbf{Z}_2$.
Rank$(\mathsf{G}_{\mathcal{L}}) = 0$, *if the correlation line* $\boldsymbol{\mathcal{L}}$ *contains the coinciding projection centers* $\mathbf{Z}_1 = \mathbf{Z}_2$.
Depending on its rank, $\mathsf{G}_{\mathcal{L}}$ *has the ranges and nullspaces shown in table 6.2.*

$\mathrm{rank}(\mathsf{G}_{\mathcal{L}})$	$\mathcal{R}(\mathsf{G}_{\mathcal{L}})$	$\mathcal{N}(\mathsf{G}_{\mathcal{L}})$	Note
2	$\boldsymbol{\eta}'' \sim \mathsf{Q}_2\boldsymbol{\mathcal{L}}$	$\boldsymbol{\eta}' \sim \mathsf{Q}_1\boldsymbol{\mathcal{L}}$	
1	a) $\mathbf{Z}_1 \in \boldsymbol{\mathcal{L}}$		
	$\mathbf{e}_1''$ (a single point)	$\{\boldsymbol{\lambda}' : \mathbf{s}'^{\top}\boldsymbol{\lambda}' = 0\}$ (a pencil of lines)	$\mathsf{G}_{\mathcal{L}} \sim \mathbf{e}_1''\mathbf{s}'^{\top}$
	b) $\mathbf{Z}_2 \in \boldsymbol{\mathcal{L}}$		
	$\mathbf{s}''$ (a single point)	$\{\boldsymbol{\lambda}' : \mathbf{e}_2'^{\top}\boldsymbol{\lambda}' = 0\}$ (a pencil of lines)	$\mathsf{G}_{\mathcal{L}} \sim \mathbf{s}''\mathbf{e}_2'^{\top}$
0	$\{\}$	$\mathbb{P}^2$	

Table 6.2: *Range* $\mathcal{R}(\mathsf{G}_{\mathcal{L}})$ *and nullspace* $\mathcal{N}(\mathsf{G}_{\mathcal{L}})$ *of the line correlation* $\mathsf{G}_{\mathcal{L}}$ *depending on its rank; cf. the figures 6.7 and 6.8. The images* Ψ_1 *resp.* Ψ_2 *are given as* $\mathsf{Q}_1 \sim \mathsf{M}_1^{-\top}[-\mathsf{S}(\mathbf{Z}_1), I_3]$ *resp.* $\mathsf{Q}_2 \sim \mathsf{M}_2^{-\top}[-\mathsf{S}(\mathbf{Z}_2), I_3]$*;* $\mathbf{e}_2'$ *resp.* $\mathbf{e}_1''$ *are their epipoles. The 3D line* $\boldsymbol{\mathcal{L}}$ *is given as* $\boldsymbol{\mathcal{L}}^{\top} \sim (\boldsymbol{L}_H{}^{\top}, \boldsymbol{L}_O{}^{\top})$*. In the two rank-1 cases the point* $\boldsymbol{s}' \sim \mathsf{M}_1\boldsymbol{L}_H$ *resp.* $\boldsymbol{s}'' \sim \mathsf{M}_2\boldsymbol{L}_H$ *is the intersection point of the line* $\boldsymbol{\mathcal{L}}$ *with the image plane* Ψ_1 *resp.* Ψ_2*.*

Proof:
First we proof the dependence of $\mathrm{rank}(\mathsf{G}_{\mathcal{L}})$ on the incidence of the projection centers $\mathbf{Z}_1$ resp. $\mathbf{Z}_2$ with the line $\boldsymbol{\mathcal{L}}$. The enclosing M matrices are assumed to be regular therefore $\mathrm{rank}(\mathsf{G}_{\mathcal{L}})$ depends only on the middle part $\overline{\mathsf{G}} = \mathsf{S}(\boldsymbol{L}_O) + \boldsymbol{Z}_2\boldsymbol{L}_H{}^{\top} - \boldsymbol{L}_H\boldsymbol{Z}_1{}^{\top}$. Without loss of generality we can assume that $\boldsymbol{\mathcal{L}}$ passes through the origin. Consequently $\boldsymbol{L}_O = \mathbf{0}$ and $\overline{\mathsf{G}} = \boldsymbol{Z}_2\boldsymbol{L}_H{}^{\top} - \boldsymbol{L}_H\boldsymbol{Z}_1{}^{\top}$. Since $\overline{\mathsf{G}}$ is made up of the sum of two dyads, all its columns are linear combinations of the two vectors $\boldsymbol{Z}_2$ and $\boldsymbol{L}_H$. Therefore $\mathrm{rank}(\mathsf{G}_{\mathcal{L}}) = 2$ in general.

Of course the rank will be 1 if $\boldsymbol{Z}_2 \sim \boldsymbol{L}_H$, which means that $\mathbf{Z}_1$ lies on $\boldsymbol{\mathcal{L}}$. The same reasoning can be applied to the rows of $\overline{\mathsf{G}}$, which shows that $\mathrm{rank}(\mathsf{G}_{\mathcal{L}}) = 1$ if $\mathbf{Z}_2$ lies on $\boldsymbol{\mathcal{L}}$. For $\overline{\mathsf{G}}$ to become $\mathsf{0}_3$ it must hold: $\boldsymbol{Z}_2\boldsymbol{L}_H{}^{\top} = \boldsymbol{L}_H\boldsymbol{Z}_1{}^{\top}$, which only can be achieved if $\boldsymbol{Z}_1 = \boldsymbol{Z}_2 = \mu\boldsymbol{L}_H$ for some scalar μ; i.e. both projection centers lie on $\boldsymbol{\mathcal{L}}$.

The next task is to prove the range and nullspaces listed in table 6.2. As an example, we will show that all mappings $\mathsf{G}_{\mathcal{L}}\boldsymbol{\lambda}'$ lie on the line $\boldsymbol{\eta}'' \sim \mathsf{Q}_2\boldsymbol{\mathcal{L}}$ in image Ψ_2; i.e. $\boldsymbol{\mathcal{L}}^{\top}\mathsf{Q}_2{}^{\top}\mathsf{G}_{\mathcal{L}}\boldsymbol{\lambda}' = 0$. With $\boldsymbol{\mathcal{L}}^{\top} \sim (\boldsymbol{L}_H{}^{\top}, \boldsymbol{L}_O{}^{\top})$ and the relation (5.1) on page 40 for the line projection matrix Q_2 we have to show $\left(\boldsymbol{L}_H{}^{\top}\mathsf{S}(\boldsymbol{Z}_2) + \boldsymbol{L}_O{}^{\top}\right)\left(\mathsf{S}(\boldsymbol{L}_O) + \boldsymbol{Z}_2\boldsymbol{L}_H{}^{\top} - \boldsymbol{L}_H\boldsymbol{Z}_1{}^{\top}\right)\mathsf{M}_1{}^{\top}\boldsymbol{\lambda}' = 0$. This condition is indeed true, which is verified by out-multiplying these two brackets and using the Plücker-identity for $\boldsymbol{\mathcal{L}}$ (3.13) and relation (B.10) for the product of two axiators. The correctness of the other range and nullspaces can be checked similarly by inserting the quantities listed in table 6.2 into the correlation proposition 6.10.

□

Because this dual correlation is always singular, there is no expression for the inverse transformation. The mapping of the lines $\boldsymbol{\lambda}''$ in image Ψ_2 to the points $\mathbf{x}'$ in image Ψ_1 using the very

same line $\mathcal{L}$ in space, however, can be represented by $\mathsf{G}_{\mathcal{L}}{}^{\top}$. This follows directly from proposition 6.10 on page 57 by exchanging the roles of the two images. In the figures 6.7 and 6.8 the two possible cases for rank-1 dual correlations are depicted; cf. table 6.2.

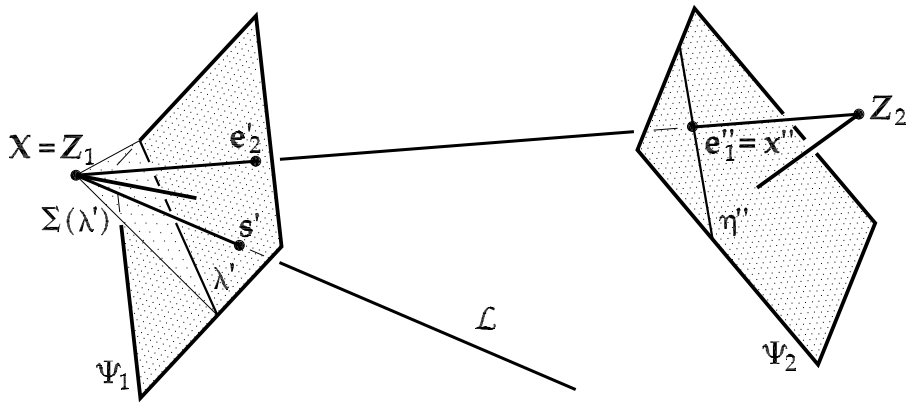

Figure 6.7: *A dual correlation from image Ψ_1 to image Ψ_2 with rank-1, where the line $\mathcal{L}$ contains the projection center $\mathbf{Z}_1$.*

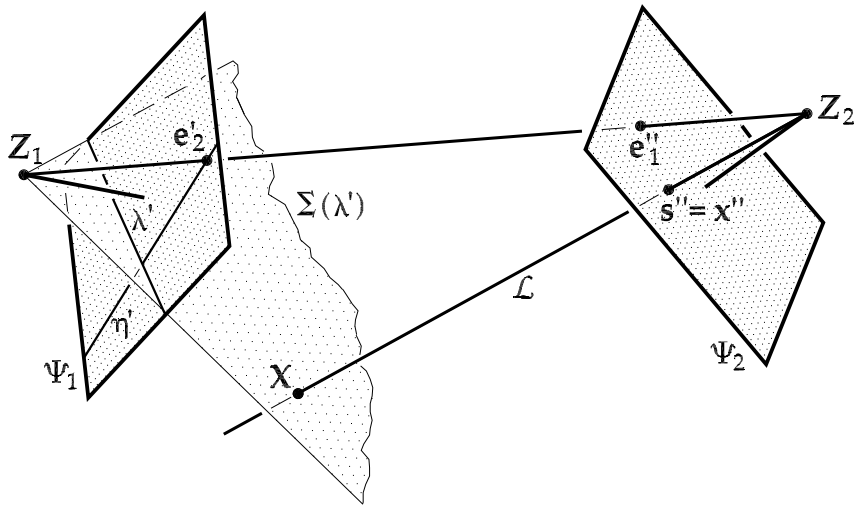

Figure 6.8: *A dual correlation from image Ψ_1 to image Ψ_2 with rank-1, where the line $\mathcal{L}$ contains the projection center $\mathbf{Z}_2$.*

For dual correlations of rank-2 the relation between the points on $\boldsymbol{\eta}'$ and the points on $\boldsymbol{\eta}''$ is one-to-one – a regular collineation in $\mathbb{P}^1$; cf. [Papadopoulo and Faugeras 1998]. This is a consequence of the following proposition.

Proposition 6.12 (Rank-2 properties of dual line correlations.) *Let a dual line correlation from image Ψ_1 to image Ψ_2 with respect to a 3D line $\mathcal{L}$ be represented by $\mathsf{G}_{\mathcal{L}}$ with $\mathrm{rank}(\mathsf{G}_{\mathcal{L}}) = 2$. Furthermore let $\mathfrak{L}_{p'}$ be the pencil of lines with carrier $\mathbf{p}'$ in image Ψ_1. If $\mathbf{p}' \in \boldsymbol{\eta}'$ then all lines $\boldsymbol{\lambda}' \not\sim \boldsymbol{\eta}'$ in $\mathfrak{L}_{p'}$ map to the very same point $\mathbf{p}''(\mathbf{p}')$ on $\boldsymbol{\eta}''$ (represented as $\mathbf{p}''(\mathbf{p}') \sim \mathsf{G}_{\mathcal{L}}\boldsymbol{\lambda}'$), with $\boldsymbol{\eta}'$ resp. $\boldsymbol{\eta}''$ being the projection of $\mathcal{L}$ into the image Ψ_1 resp. Ψ_2; i.e. the right and left nullspace of $\mathsf{G}_{\mathcal{L}}$. The point $\mathbf{p}''(\mathbf{p}')$ is the map of the intersection point of $\mathcal{L}$ with the projection ray of $\mathbf{p}'$.*

If $\mathbf{p}' \notin \boldsymbol{\eta}'$ then each line $\boldsymbol{\lambda}'$ in $\mathfrak{L}_{p'}$ is mapped to a different point on $\boldsymbol{\eta}''$.

Proof:
The proof is easy, however a bit too long to be included in detail. The basic idea is to represent the lines $\boldsymbol{\lambda}'$ of the pencil $\mathfrak{L}_{p'}$ as $\boldsymbol{\lambda}' \sim \mathsf{S}(\mathbf{p}')\mathbf{q}'$, with $\mathbf{q}'$ being an arbitrary point $\neq \mathbf{p}'$ and to express $\mathbf{q}'$ as a linear combination of three linearly independent vectors as

$$\mathbf{q}' \sim a\mathbf{p}' + b\mathsf{M}_1\boldsymbol{L}_H + c\mathbf{p}' \times (\mathsf{M}_1\boldsymbol{L}_H).$$

with the correlation line $\boldsymbol{\mathcal{L}}$ given as $\boldsymbol{\mathcal{L}}^\top \sim (\boldsymbol{L}_H{}^\top, \boldsymbol{L}_O{}^\top)$. Again as in the previous proof we can assume without loss of generality that $\boldsymbol{\mathcal{L}}$ passes through the origin and so $\boldsymbol{L}_O = \mathbf{0}$. With this preparation and proposition 6.10 on page 57 we can write

$$\mathbf{p}''(\mathbf{p}') \sim \mathsf{G}_{\mathcal{L}}\boldsymbol{\lambda}' \sim \mathsf{M}_2\left(\boldsymbol{Z}_2\boldsymbol{L}_H{}^\top - \boldsymbol{L}_H\boldsymbol{Z}_1{}^\top\right)\mathsf{M}_1{}^\top\mathsf{S}(\mathbf{p}')\left[\mathsf{M}_1\boldsymbol{L}_H, \mathbf{p}' \times (\mathsf{M}_1\boldsymbol{L}_H)\right]\begin{pmatrix} b \\ c \end{pmatrix},$$

which does not depend on the coefficient a. Now we define $\mathbf{p}'$ to lie on $\boldsymbol{\eta}'$, thus the projection ray of $\mathbf{p}'$ is coplanar with the line $\boldsymbol{\mathcal{L}}$. Using equation (5.23) on page 42 for the projection ray and the coplanarity condition for lines (3.16) on page 20 we get the following constraint

$$\boldsymbol{L}_H^\top\mathsf{S}(\boldsymbol{Z}_1)\mathsf{M}_1{}^{-1}\mathbf{p}' = 0.$$

Considering this constraint in the above expression for $\mathsf{G}_{\mathcal{L}}\boldsymbol{\lambda}'$, we see that $\mathbf{p}''(\mathbf{p}')$ is also independent of the coefficient b and only c remains, which then only acts as a scale.

Therefore $\mathsf{G}_{\mathcal{L}}\boldsymbol{\lambda}'$ is independent of $\mathbf{q}'$ and so all lines $\boldsymbol{\lambda}'$ of the pencil $\mathfrak{L}_{p'}$ are mapped to the very same point $\mathbf{p}''(\mathbf{p}')$ in image Ψ_2. If the line $\boldsymbol{\lambda}' \sim \boldsymbol{\eta}'$, then $\mathbf{q}' \in \boldsymbol{\eta}'$ and $c = 0$, and therefore $\mathbf{p}''(\mathbf{p}') = \mathbf{0}$.

Note that it was assumed that $\mathbf{p}'$ and $\mathsf{M}_1\boldsymbol{L}_H$ are linearly independent. If this assumption is not valid (i.e. $\mathbf{p}' \sim \mathsf{M}_1\boldsymbol{L}_H$) it is straightforward to show that $\mathbf{p}''(\mathbf{p}')$ is still independent of $\mathbf{q}'$. This shows the one-to-one relation by $\mathsf{G}_{\mathcal{L}}$ between the points on $\boldsymbol{\eta}'$ and the points on $\boldsymbol{\eta}''$.

Conversely if $\mathbf{p}'$ does not lie on $\boldsymbol{\eta}'$ the dependence on $\mathbf{q}'$ remains and each line $\boldsymbol{\lambda}' \sim \mathsf{S}(\mathbf{p}')\mathbf{q}'$ of the pencil is mapped to a different point on $\boldsymbol{\eta}''$ as the relation $\mathsf{G}_{\mathcal{L}}\boldsymbol{\lambda}'$ can then be expressed as $\mathsf{H}_{\mathcal{L},p'}\mathbf{q}'$, with the rank-2 homography $\mathsf{H}_{\mathcal{L},p'} \sim \mathsf{G}_{\mathcal{L}}\mathsf{S}(\mathbf{p}')$.

□

In contrast to the plane homographies of the previous section, where the homography $\mathsf{H}_\infty \sim \mathsf{M}_2\mathsf{M}_1{}^{-1}$ induced by the plane at infinity played a special role, no special dual line correlation exists. Note that the expression $\mathsf{M}_2\mathsf{M}_1^\top$ is regular and therefore does not represent a dual line correlation.

Chapter 7

Three-View Geometry – the Trifocal Tensor

In section 6.1.1 we saw that corresponding points in two images have to fulfil the epipolar constraint, which is algebraically represented as the homogenous bilinear form (6.2) whose coefficients are made up of the elements of the fundamental matrix. Motivated by this observation we can ask for similar *constraints which must hold between corresponding features in three images*. If we consider especially a triple of corresponding lines we will get two such constraints as it is shown in section 7.3. These constraints are represented as homogenous trilinear forms whose coefficients are made up by the elements of a valence-3 tensor – *the trifocal tensor*, which is derived in the following section.

7.1 A derivation of the trifocal tensor

We consider three images Ψ_h $(h = 1, 2, 3)$ represented by their point projection matrices $\mathsf{P}_h \sim \mathsf{M}_h\,[I_3, -\boldsymbol{Z}_h]$ with regular M_h, together with a corresponding triple of image lines $\boldsymbol{\lambda}_h$, which are the mappings of a 3D line $\boldsymbol{\mathcal{L}}$; see figure 7.1 on the next page. In general $\boldsymbol{\mathcal{L}}$ is uniquely determined by the intersection of the back projection of two of these image lines (e.g. $\boldsymbol{\lambda}''$ and $\boldsymbol{\lambda}'''$). Let $\overline{\boldsymbol{\mathcal{L}}}$ denote this intersection of the projection planes $\boldsymbol{\Sigma}(\boldsymbol{\lambda}'')$ and $\boldsymbol{\Sigma}(\boldsymbol{\lambda}''')$. Then we will obtain two constraints in demanding that the map of $\overline{\boldsymbol{\mathcal{L}}}$ into image Ψ_1 is identical to $\boldsymbol{\lambda}'$.

Algebraically the three image lines $\boldsymbol{\lambda}_h$ are represented as homogenous vectors $\boldsymbol{\lambda}_h$. The projection planes $\boldsymbol{\Sigma}(\boldsymbol{\lambda}'')$ and $\boldsymbol{\Sigma}(\boldsymbol{\lambda}''')$ are represented as $\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Sigma}_3$. Using the back projection of lines (5.20) these planes are given as $\boldsymbol{\Sigma}_2 \sim \mathsf{P}_2{}^\top \boldsymbol{\lambda}''$ and $\boldsymbol{\Sigma}_3 \sim \mathsf{P}_3{}^\top \boldsymbol{\lambda}'''$. Using the relation for the intersection of two planes (3.30) the line $\overline{\boldsymbol{\mathcal{L}}}$ is represented as

$$\overline{\boldsymbol{\mathcal{L}}} \sim \Pi^*(\boldsymbol{\Sigma}_2)\boldsymbol{\Sigma}_3 = \begin{bmatrix} -\mathsf{S}(\mathsf{M}_2{}^\top \boldsymbol{\lambda}'') & \mathbf{0} \\ (\boldsymbol{Z}_2{}^\top \mathsf{M}_2{}^\top \boldsymbol{\lambda}'') I_3 & \mathsf{M}_2{}^\top \boldsymbol{\lambda}'' \end{bmatrix} \begin{bmatrix} \mathsf{M}_3{}^\top \\ -\boldsymbol{Z}_3{}^\top \mathsf{M}_3{}^\top \end{bmatrix} \boldsymbol{\lambda}'''.$$

The projection of $\overline{\boldsymbol{\mathcal{L}}}$ into image Ψ_1 must be identical to $\boldsymbol{\lambda}'$. Using the line projection matrix Q_1 from proposition 5.1 this can be represented as

$$\begin{aligned} \boldsymbol{\lambda}' &\sim \mathsf{Q}_1 \overline{\boldsymbol{\mathcal{L}}} \sim \mathsf{M}_1{}^{-\top} [-\mathsf{S}(\boldsymbol{Z}_1), I_3]\, \overline{\boldsymbol{\mathcal{L}}} \\ &\sim \mathsf{M}_1{}^{-\top} \mathsf{S}(\boldsymbol{Z}_1) \mathsf{S}(\mathsf{M}_2{}^\top \boldsymbol{\lambda}'') \mathsf{M}_3{}^\top \boldsymbol{\lambda}''' + \mathsf{M}_1{}^{-\top} (\mathsf{M}_2{}^\top \boldsymbol{Z}_2{}^\top \boldsymbol{\lambda}'') \mathsf{M}_3{}^\top \boldsymbol{\lambda}''' - \mathsf{M}_1{}^{-\top} \mathsf{M}_2{}^\top \boldsymbol{\lambda}'' (\boldsymbol{Z}_3{}^\top \mathsf{M}_3{}^\top \boldsymbol{\lambda}'''). \end{aligned}$$

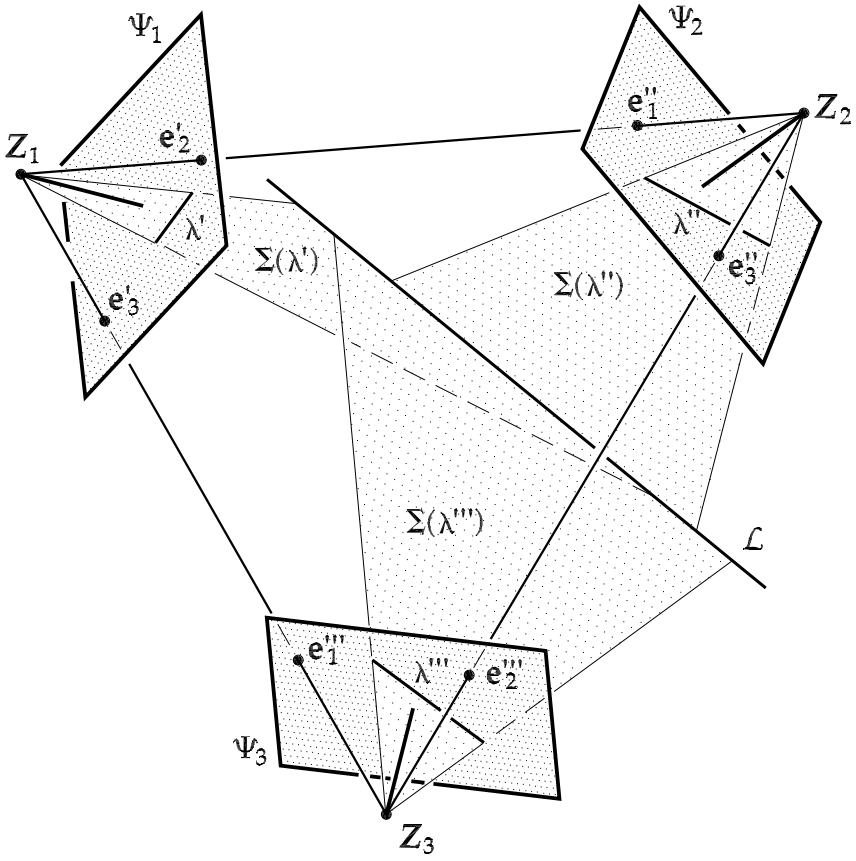

Figure 7.1: *Three images $\{\Psi_1, \Psi_2, \Psi_3\}$, a corresponding triple of image lines $\{\boldsymbol{\lambda}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''\}$ and the respective line in space $\boldsymbol{\mathcal{L}}$.*

With the axiator relation (B.10) we can rewrite $\mathsf{S}(\boldsymbol{Z}_1)\mathsf{S}(\mathsf{M}_2{}^\top\boldsymbol{\lambda}'')$ as $\mathsf{M}_2{}^\top\boldsymbol{\lambda}''\boldsymbol{Z}_1{}^\top - \boldsymbol{Z}_1{}^\top\mathsf{M}_2{}^\top\boldsymbol{\lambda}''\boldsymbol{I}_3$ and get

$$\boldsymbol{\lambda}' \sim -((\mathsf{M}_2\boldsymbol{Z}_1 - \mathsf{M}_2\boldsymbol{Z}_2)^\top\boldsymbol{\lambda}'')\mathsf{M}_1{}^{-\top}\mathsf{M}_3{}^\top\boldsymbol{\lambda}''' + ((\mathsf{M}_3\boldsymbol{Z}_1 - \mathsf{M}_3\boldsymbol{Z}_3)^\top\boldsymbol{\lambda}''')\mathsf{M}_1{}^{-\top}\mathsf{M}_2{}^\top\boldsymbol{\lambda}''. \tag{7.1}$$

Now we introduce the epipoles $\mathbf{e}''_1$ and $\mathbf{e}'''_1$, cf. epipole equation (6.1) on page 45,

$$\begin{aligned}\mathbf{e}''_1 &= \mathsf{M}_2(\boldsymbol{Z}_1 - \boldsymbol{Z}_2)\\ \mathbf{e}'''_1 &= \mathsf{M}_3(\boldsymbol{Z}_1 - \boldsymbol{Z}_3),\end{aligned} \tag{7.2}$$

and although they are represented as projective vectors, their scale is not independent from the other orientation elements of the three images. This is emphasized by using '=' instead of '$\sim$'. We also give the definition of the other four possible epipoles in the three images as

$$\begin{aligned}\mathbf{e}'_2 &= \mathsf{M}_1(\boldsymbol{Z}_2 - \boldsymbol{Z}_1)\\ \mathbf{e}'_3 &= \mathsf{M}_1(\boldsymbol{Z}_3 - \boldsymbol{Z}_1)\\ \mathbf{e}''_3 &= \mathsf{M}_2(\boldsymbol{Z}_3 - \boldsymbol{Z}_2)\\ \mathbf{e}'''_2 &= \mathsf{M}_3(\boldsymbol{Z}_2 - \boldsymbol{Z}_3).\end{aligned} \tag{7.3}$$

Finally we introduce the matrices A and B as follows

$$\begin{aligned} \mathsf{A} &= \mathsf{M}_2\mathsf{M}_1^{-1} \\ \mathsf{B} &= \mathsf{M}_3\mathsf{M}_1^{-1}. \end{aligned} \tag{7.4}$$

Observe that A and B represent the homographies from image Ψ_1 to the other two images induced by the plane at infinity (cf. equation 6.17 on page 56). With these new quantities we write equation (7.1) as

$$\boldsymbol{\lambda}' \sim -(\mathbf{e}_1''^{\top}\boldsymbol{\lambda}'')\mathsf{B}^{\top}\boldsymbol{\lambda}''' + (\mathbf{e}_1'''^{\top}\boldsymbol{\lambda}''')\mathsf{A}^{\top}\boldsymbol{\lambda}''. \tag{7.5}$$

This relation expresses the line $\boldsymbol{\lambda}'$ as a bilinear function in $\boldsymbol{\lambda}''$ and $\boldsymbol{\lambda}'''$ and is given in matrix notation. Because of the multiplication rule associated with this notation (*multiply left row with right column*) the relation (7.5) can not be further simplified.

However, using tensor notation we can reach a more simpler representation. The epipoles $\mathbf{e}_1''$ and $\mathbf{e}_1'''$ are contra-variant coordinate vectors which are indexed by a super-script. The column vectors $\boldsymbol{\lambda}'$, $\boldsymbol{\lambda}''$ and $\boldsymbol{\lambda}'''$ are vectors of the dual space and therefore co-variant coordinate vectors, which are indexed by a sub-script. The matrices A and B require two indices – one sub-script and one super-script. This follows from the summation convention and the fact that both matrices represent collineations; cf. section 4. The first index corresponds to the rows and the second one to the columns of the matrix. Therefore we have the following transition between matrix and tensor notation

$$\boldsymbol{\lambda}' = \lambda_i' \qquad \boldsymbol{\lambda}'' = \lambda_j'' \qquad \boldsymbol{\lambda}''' = \lambda_k'''$$

$$\mathbf{e}_1'' = e_1''^{j} \qquad \mathbf{e}_1''' = e_1'''^{k}$$

$$\mathsf{A}^{\top} = (A^{\top})_{i\cdot}^{\ j} = A_{\cdot i}^{j} \qquad \mathsf{B}^{\top} = (B^{\top})_{i\cdot}^{\ k} = B_{\cdot i}^{k}$$

with $i, j, k = 1, 2, 3$. And so relation (7.5) is given in tensor notation as

$$\lambda_i' \sim -e_1''^{j}\lambda_j''B_{\cdot i}^{k}\lambda_k''' + e_1'''^{k}\lambda_k'''A_{\cdot i}^{j}\lambda_j'' \sim \lambda_j''\lambda_k'''T_{\cdot\cdot i}^{kj}. \tag{7.6}$$

By combining the quantities that depend on the orientation parameters of the three images we found the *trifocal tensor* $T_{\cdot\cdot i}^{kj}$.

Definition 7.1 *Any homogenous* valence-*(1,2) tensor $T_{\cdot\cdot i}^{kj}$, that can be represented as*

$$T_{\cdot\cdot i}^{kj} \sim e_1''^{j}B_{\cdot i}^{k} - e_1'''^{k}A_{\cdot i}^{j}, \tag{7.7}$$

using two matrices A and B, and two vectors $\mathbf{e}_1''$ and $\mathbf{e}_1'''$, is a ***trifocal tensor****.*

Note, however, that the representation of a given trifocal tensor in $\mathbf{e}_1''$, $\mathbf{e}_1'''$, A and B is not unique, because in section 7.4.3 on page 90 we will see that the same tensor can also be represented using $w\mathbf{e}_1''$, $w\mathbf{e}_1'''$, $\mathsf{A} + \mathbf{e}_1''\mathbf{v}^{\top}$ and $\mathsf{B} + \mathbf{e}_1'''\mathbf{v}^{\top}$ for any arbitrary vector $\mathbf{v}$ and $w \neq 0$.

Because the trifocal tensor is a homogenous valence-(1,2) tensor it has $3^3 = 27$ elements. It only depends on the *relative* orientation of the three images. This can be seen by its defining quantities. The epipoles ($\mathbf{e}_1''$, $\mathbf{e}_1'''$) and the point-homographies (A, B) do not depend on the

choice of the object system. Therefore the trifocal tensor plays the same role for three images as the fundamental matrix plays for two, e.g. the trifocal tensor contains all the information of the epipolar geometry of three images and can be determined by image observations alone – including both lines *and* points as we will see later. Lines can not be used for the fundamental matrix, since the projection planes of two image lines intersect in any case; i.e. even if the two lines are not corresponding.

Observe that one image plays a special role in this derivation. It is the image in which the 3D line $\overline{\mathcal{L}}$ is projected using the respective Q matrix (in this case image Ψ_1). The index of the special image appears as the sub-script in the trifocal tensor $T^{kj}_{\cdot\cdot i}$. Since we can choose any of the three given images to play that special role and can switch the roles of the other two, there are *six different* trifocal tensors associated with three given images: $T(\Psi_1, \Psi_2, \Psi_3)$, $T(\Psi_1, \Psi_3, \Psi_2)$, $T(\Psi_2, \Psi_1, \Psi_3)$, $T(\Psi_2, \Psi_3, \Psi_1)$, $T(\Psi_3, \Psi_1, \Psi_2)$, $T(\Psi_3, \Psi_2, \Psi_1)$. For any given trifocal tensor the other 5 representations can be derived from the given one; see section 7.5.

Although the trifocal tensor is made up of 27 elements it only has 18 degrees of freedom in case of uncalibrated images, which follows from proposition 6.2 on page 44. Therefore the tensor elements have to fulfil 9 constraints, one of which is the fixing of the scale, since the trifocal tensor is a homogenous quantity.

Since the trifocal tensor is indexed by three indices, we can imagine the tensor as a $3\times3\times3$ cube of numbers. If we fix one of the three indices, only two free indices remain. We slice – so to say – a matrix out of the cube. Since we have three indices to fix, we get three different kinds of slices. These slices, which will be investigated in detail in section 7.2, are the basic input for finding the constraints for the trifocal tensor. So actually we will not work much with the trifocal tensor in its three-index-representation (7.7), but much more with those matrices – the *tensorial slices*.

A few words about the placement of the indices. The placement of the super and sub-scripts in the tensor equation (7.7) is unusual compared to other authors, who refer to the trifocal tensor normally as T^{jk}_i; e.g. [Hartley and Zisserman 2001] or [Heyden 2000]. The different placement is caused by different approaches for the derivation of the trifocal tensor. The algebraic derivation in [Heyden 2000], which is also summarized in section C, does not enforce a special order of the indices. The geometric based derivation in matrix notation used in this section, however, enforces the order of the indices shown in equation (7.7). In [Hartley and Zisserman 2001] also the same geometric derivation based on matrix notation is presented. Their transition to tensor notation, however, is is done disadvantageously as the rows of a matrix are associated with contra-variant indices and the columns are associated with co-variant indices – a method that would prevent a tensorial representation of the fundamental matrix, which requires two sub-scripts. In this thesis we use the usual method in tensor calculus, where the order (indicated with the dots $\cdot$) determines the row and column indices. As long as one works only in tensor notation and sticks with the chosen placement of the indices no problems arise. In this thesis, however, we will actually not work very much with the tensor in its representation with the three indices, but more with the above mentioned matrices (the tensorial slices). Therefore it is required to have a rigid transition between the tensor notation and the matrix notation of these slices.

If we had chosen to represent the matrices $\mathsf{M}_2\mathsf{M}_1^{-1}$ and $\mathsf{M}_3\mathsf{M}_1^{-1}$ in equation set (7.4) with $\mathsf{A}^\top$ and $\mathsf{B}^\top$ respectively, we would have obtained a tensor representation as $T_{i\cdot\cdot}^{\ jk}$. However, as

these matrices represent *point* homographies, it is not advisable to use the notation of a *line* homography.

7.1.1 Computing the trifocal tensor from given projection matrices

Assume we are given three images Ψ_1, Ψ_2 and Ψ_3 with their projection matrices

$$\begin{aligned}
\mathsf{P}_1 &\sim \mathsf{M}_1\,[I_3, -\boldsymbol{Z}_1] \sim \begin{bmatrix} \boldsymbol{\Pi}_1'^{\top} \\ \boldsymbol{\Pi}_2'^{\top} \\ \boldsymbol{\Pi}_3'^{\top} \end{bmatrix} \\
\mathsf{P}_2 &\sim \mathsf{M}_2\,[I_3, -\boldsymbol{Z}_2] \sim \begin{bmatrix} \boldsymbol{\Pi}_1''^{\top} \\ \boldsymbol{\Pi}_2''^{\top} \\ \boldsymbol{\Pi}_3''^{\top} \end{bmatrix} \\
\mathsf{P}_3 &\sim \mathsf{M}_3\,[I_3, -\boldsymbol{Z}_3] \sim \begin{bmatrix} \boldsymbol{\Pi}_1'''^{\top} \\ \boldsymbol{\Pi}_2'''^{\top} \\ \boldsymbol{\Pi}_3'''^{\top} \end{bmatrix}.
\end{aligned}$$

With these projection matrices we can determine the trifocal tensor of the three images. If the M matrices are regular we can use the relations obtained in the previous section where we derived the trifocal tensor:

$$\begin{aligned}
\mathbf{e}_1'' = \mathsf{M}_2(\boldsymbol{Z}_1 - \boldsymbol{Z}_2) \quad &\rightarrow \quad e_1''^{j} \\
\mathbf{e}_1''' = \mathsf{M}_3(\boldsymbol{Z}_1 - \boldsymbol{Z}_3) \quad &\rightarrow \quad e_1'''^{j} \\
\mathsf{A} = \mathsf{M}_2\mathsf{M}_1^{-1} \quad &\rightarrow \quad A_{i.}^{\ j} \\
\mathsf{B} = \mathsf{M}_3\mathsf{M}_1^{-1} \quad &\rightarrow \quad B_{i.}^{\ k} \\
&\Rightarrow \quad T_{..i}^{kj} \sim e_1''^{j} B_{.i}^{k} - e_1'''^{k} A_{.i}^{j}.
\end{aligned} \tag{7.8}$$

For the derivation in the previous section we required the three M matrices to be regular; i.e. the projection centers must not lie in the plane at infinity $\boldsymbol{\Pi}_\infty$. If this requirement is not fulfilled, then the trifocal tensor can be computed using the following relation, which is obtained from a more algebraic approach in deriving the trifocal tensor (equation (C.5) on page 186):

$$T_{..i}^{kj} = \frac{1}{2}\epsilon_{ief}\ \det \begin{bmatrix} \boldsymbol{\Pi}_e'^{\top} \\ \boldsymbol{\Pi}_f'^{\top} \\ \boldsymbol{\Pi}_j''^{\top} \\ \boldsymbol{\Pi}_k'''^{\top} \end{bmatrix}, \tag{7.9}$$

where ϵ_{ief} is the fully antisymmetric valence-3 tensor introduced in section 4. Note, that in (7.9) it is summed over e and f.

7.2 The tensorial slices

Because the trifocal tensor is indexed by three indices, each running from 1 to 3, we can imagine the tensor as a $3\times3\times3$ cube of numbers. If we fix one of the indices i, j, k at a certain value (1,2 or 3), we slice – so to say – a 3×3 matrix out of the tensor; see figure 7.2. Since we have three indices to choose, we can slice in three different directions and thus get three different types of matrices, which we will call conveniently I_i, J_j and K_k. And since in each direction the fixed index can take three values, we get in total 9 possible slices out of the tensor.

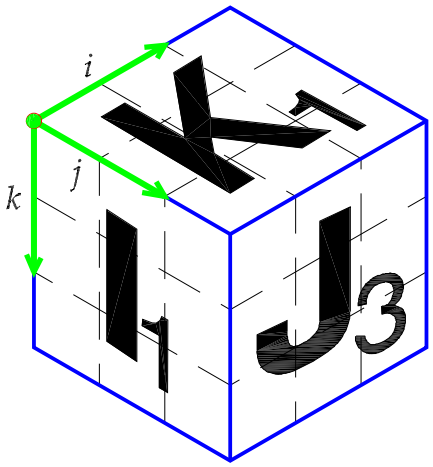

Figure 7.2: *The trifocal tensor interpreted as a* $3\times3\times3$ *cube of numbers with the tensorial slices* I_1, J_3 *and* K_1.

7.2.1 The correlation slices

Let us start by fixing the i-index in the expression (7.7) for the tensor. This returns a two-fold indexed quantity $(I_i)^{kj}$

$$(I_i)^{kj} = T^{kj}_{\cdot\,\cdot\,(i)} = e_1''^{\,j} B^k_{\cdot(i)} - e_1'''^{\,k} A^j_{\cdot(i)}.$$

Putting an index in brackets is the usual way in tensor notation to indicate the fixation of this index. Since $(I_i)^{kj}$ is a two-fold indexed quantity, we can interpret it as a 3×3 matrix I_i. Its rows are created by the first index (k) and its columns by the second index (j). Therefore $A^j_{\cdot(i)}$ corresponds to $\mathsf{A}\,\boldsymbol{i}_i$ and $B^k_{\cdot(i)}$ to $\mathsf{B}\,\boldsymbol{i}_i$, where $\boldsymbol{i}_i$ is the ith column vector of the identity matrix $\boldsymbol{I}_3$. And so we get I_i in matrix notation as

$$\mathsf{I}_i = \mathsf{B}\,\boldsymbol{i}_i\,\mathbf{e}_1''^{\top} - \mathbf{e}_1'''\,\boldsymbol{i}_i{}^{\top}\mathsf{A}^{\top}. \tag{7.10}$$

Note that $\mathsf{A}\,\boldsymbol{i}_i$ and $\mathbf{e}_1''$ need to be transposed because the j-index counts the columns in I_i. We can further express I_i in terms of the projection matrices $\mathsf{P}_h \sim \mathsf{M}_h\,[\boldsymbol{I}_3, -\boldsymbol{Z}_h]$, $(h = 1,2,3)$, with regular M_h using the relations (7.4) and (7.2) as

$$\mathsf{I}_i = \mathsf{M}_3\Big(\underbrace{-\boldsymbol{Z}_1\boldsymbol{i}_i{}^{\top}\mathsf{M}_1{}^{-\top} + \mathsf{M}_1^{-1}\boldsymbol{i}_i\boldsymbol{Z}_1{}^{\top}}_{\mathsf{S}(\boldsymbol{R}'_{i,O})} + \boldsymbol{Z}_3\underbrace{\boldsymbol{i}_i{}^{\top}\mathsf{M}_1{}^{-\top}}_{\boldsymbol{R}'_{i,H}{}^{\top}} - \underbrace{\mathsf{M}_1{}^{-1}\boldsymbol{i}_i}_{\boldsymbol{R}'_{1i,H}}\boldsymbol{Z}_2{}^{\top}\Big)\mathsf{M}_2{}^{\top}, \tag{7.11}$$

where $\boldsymbol{\mathcal{R}}'_i \sim (\boldsymbol{R}'^{\top}_{i,H}, \boldsymbol{R}'^{\top}_{i,O})$ is the ith principal ray of image Ψ_1; see relation (5.18) on page 41. The comparison of the equation (7.11) with proposition 6.10 on page 57 reveals that l_i describes a *dual correlation* of the lines in image Ψ_2 to the points in image Ψ_3 via the ith principal ray $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1; see figure 7.3. For this reason we will refer to these slices as **correlation slices**[1] and drop the "dual" to simplify matters. The following proposition summarizes these findings.

Proposition 7.1 (The correlation slices l_i of the trifocal tensor.) *Let the trifocal tensor be given as $T^{kj}_{\cdot\cdot i} \sim e''^{j}_1 B^k_{\cdot i} - e'''^{k}_1 A^j_{\cdot i}$. Then the correlation slices l_i, which are obtained by fixing the i-index, are represented as $\mathsf{l}_i = \mathsf{B}\ \boldsymbol{i}_i\ \mathbf{e}''^{\top}_1 - \mathbf{e}'''_1\ \boldsymbol{i}_i{}^{\top}\mathsf{A}^{\top}$. The matrix l_i represents a dual correlation of the lines in image Ψ_2 to the points in image Ψ_3 via the ith principal ray $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1. The matrix $\mathsf{l}_i{}^{\top}$ maps the lines in image Ψ_3 to the points in image Ψ_2 via the very same principal ray.*

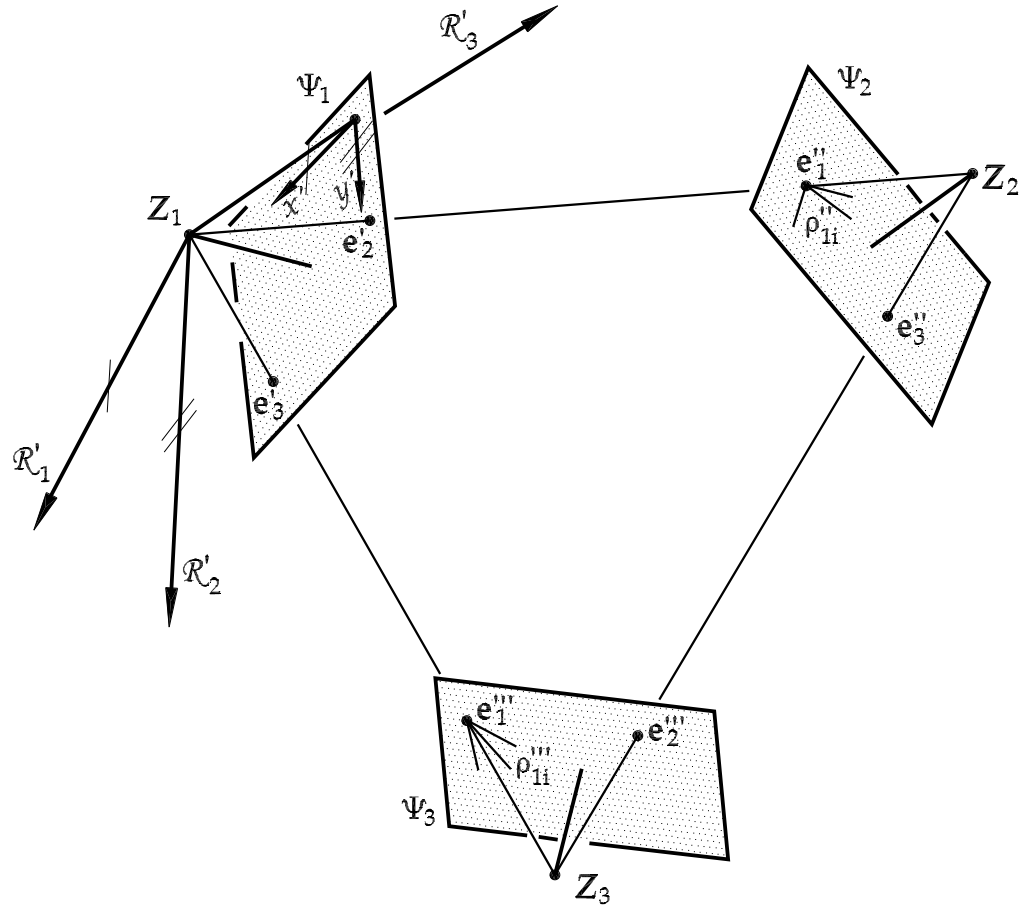

Figure 7.3: *The l_i matrices represent dual line correlations from image Ψ_2 to image Ψ_3 induced by the principal rays $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1. The maps of these principal rays in the images Ψ_2 and Ψ_3 are the lines $\boldsymbol{\rho}''_{1i}$ and $\boldsymbol{\rho}'''_{1i}$ respectively.*

Proposition 7.2 (Linear combinations of the correlation slices l_i.) *The matrix $\mathsf{l}_\diamond(\mathbf{x}')$, which is the linear combination of the homography slices l_i given by*

$$\mathsf{l}_\diamond(\mathbf{x}') = u'\mathsf{l}_1 + v'\mathsf{l}_2 + w'\mathsf{l}_3, \tag{7.12}$$

[1] In [Faugeras and Luong 2001] the respective slices are called *trifocal matrices* G_i. Due to their choice of indices the matrices G_i correspond to the matrices $\mathsf{l}_i{}^{\top}$ considered in this thesis.

or

$$\mathsf{l}_{\diamond}(\mathbf{x}') = \mathsf{M}_3\Big(-\boldsymbol{Z}_1\mathbf{x}'^{\top}\mathsf{M}_1{}^{-\top} + \mathsf{M}_1^{-1}\mathbf{x}'\boldsymbol{Z}_1{}^{\top} + \boldsymbol{Z}_3\mathbf{x}'^{\top}\mathsf{M}_1{}^{-\top} - \mathsf{M}_1{}^{-1}\mathbf{x}'\boldsymbol{Z}_2{}^{\top}\Big)\mathsf{M}_2{}^{\top},$$

describes the dual correlation of the lines in image Ψ_2 to the points in image Ψ_3 via the projection ray of the point $\mathbf{x}' \sim (u', v', w')^{\top}$ in image Ψ_1.

Proof:
The comparison of the second expression for $\mathsf{l}_{\diamond}(\mathbf{x}')$ with relation (7.11) shows that $\mathsf{l}_{\diamond}(\mathbf{x}')$ corresponds to the dual correlation matrix with respect to a 3D line passing through $\mathbf{Z}_1$ and whose direction is given by $\mathsf{M}_1{}^{-1}\mathbf{x}'$; i.e. the projection ray of the point $\mathbf{x}'$ in the image Ψ_1.

□

Adapting from proposition 6.11 on page 57 we can give the rank of l_i depending on the arrangement of the images Ψ_2 and Ψ_3 with respect to the image Ψ_1.

Proposition 7.3 (The ranks of the correlation slices l_i.) *In general the rank of l_i) is 2. The right ($\boldsymbol{\rho}''_{1i}$ resp. left ($\boldsymbol{\rho}'''_{1i}$) nullspace of l_i is the map of the* ith *principal ray $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1 into image Ψ_2 resp. Ψ_3.*
The rank of l_i is 1, if the ith *principal ray $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1 contains at least one of the projection centers of the other two images, with $\mathbf{Z}_2 \neq \mathbf{Z}_3$. If $\mathbf{Z}_2 \in \boldsymbol{\mathcal{R}}'_i$ then $\mathsf{l}_i \sim \mathbf{e}'''_2\mathbf{e}''^{\top}_1$. If $\mathbf{Z}_3 \in \boldsymbol{\mathcal{R}}'_i$ then $\mathsf{l}_i \sim \mathbf{e}'''_1\mathbf{e}''^{\top}_3$.*
The rank of l_i is 0, if the ith *principal ray $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1 contains the coinciding projection centers $\mathbf{Z}_2 = \mathbf{Z}_3$.*
Depending on its rank, l_i has the ranges and nullspaces shown in table 7.1.

rank(l_i)	$\mathcal{R}(\mathsf{l}_i)$	$\mathcal{N}(\mathsf{l}_i)$	Note
2	$\boldsymbol{\rho}'''_{1i} \sim \mathsf{Q}_3\boldsymbol{\mathcal{R}}'_i$	$\boldsymbol{\rho}''_{1i} \sim \mathsf{Q}_2\boldsymbol{\mathcal{R}}'_i$	
1	a) $\mathbf{Z}_2 \in \boldsymbol{\mathcal{R}}'_i$		
	$\mathbf{e}'''_2$ (a single point)	$\{\boldsymbol{\lambda}'' : \mathbf{e}''^{\top}_1\boldsymbol{\lambda}'' = 0\}$ (a pencil of lines)	$\mathsf{l}_i \sim \mathbf{e}'''_2\mathbf{e}''^{\top}_1$
	b) $\mathbf{Z}_3 \in \boldsymbol{\mathcal{R}}'_i$		
	$\mathbf{e}'''_1$ (a single point)	$\{\boldsymbol{\lambda}'' : \mathbf{e}''^{\top}_3\boldsymbol{\lambda}'' = 0\}$ (a pencil of lines)	$\mathsf{l}_i \sim \mathbf{e}'''_1\mathbf{e}''^{\top}_3$
0	$\{\}$	$\mathbb{P}^2$	

Table 7.1: *Range $\mathcal{R}(\mathsf{l}_i)$ and nullspace $\mathcal{N}(\mathsf{l}_i)$ of the correlation slice l_i depending on its rank. The images Ψ_2 resp. Ψ_3 are given as $\mathsf{Q}_2 \sim \mathsf{M}_2{}^{-\top}\left[-\mathsf{S}(\mathbf{Z}_2), \mathit{l}_3\right]$ resp. $\mathsf{Q}_3 \sim \mathsf{M}_3{}^{-\top}\left[-\mathsf{S}(\mathbf{Z}_3), \mathit{l}_3\right]$; $\mathbf{e}''_1$ and $\mathbf{e}''_3$ are the epipoles in image Ψ_2 and $\mathbf{e}'''_1$ and $\mathbf{e}'''_2$ are the epipoles in image Ψ_3. The* ith *principal ray of image Ψ_1 is given as $\boldsymbol{\mathcal{R}}'_i$. Note that all these properties hold also for $\mathsf{l}_{\diamond}(\mathbf{x}') \sim x'^i\mathsf{l}_i$, but by exchanging $\boldsymbol{\mathcal{R}}'_i$ with $\boldsymbol{\mathcal{S}}(\mathbf{x}')$ – the projection ray through the point $\mathbf{x}'$ in image Ψ_1.*

For rank-2 correlation slices we can find the following important properties; [Spetsakis and Aloimonos 1990] or [Papadopoulo and Faugeras 1998]. We will use them in section 7.6 for deriving the internal constraints of the trifocal tensor and in section 7.4.1 for determining the epipoles $\boldsymbol{e}''_1$ and $\boldsymbol{e}'''_1$.

Proposition 7.4 (Properties of rank-2 correlation slices.) *If the correlation slices* l_i *have rank-2, then the following holds:*

(a) The right and left nullspaces $\mathcal{N}(\mathsf{l}_i)$ *and* $\mathcal{N}(\mathsf{l}_i^\top)$ *of* l_i *are the maps* $\boldsymbol{\rho}''_{1i}$ *and* $\boldsymbol{\rho}'''_{1i}$ *of the* i*th principal ray of image* Ψ_1 *into image* Ψ_2 *and* Ψ_3 *respectively; cf. figure 7.3 on page 67. In other words: the lines* $\boldsymbol{\rho}''_{1i}$ *and* $\boldsymbol{\rho}'''_{1i}$ *are the epipolar lines in image* Ψ_2 *and* Ψ_3 *corresponding to the canonical points of image* Ψ_1*; i.e.* $\boldsymbol{\rho}''_{1i} \sim \mathsf{F}_{12}{}^\top \boldsymbol{i}_i$ *and* $\boldsymbol{\rho}'''_{1i} \sim \mathsf{F}_{13}{}^\top \boldsymbol{i}_i$*, with* F_{12} *and* F_{13} *being the respective fundamental matrices.*

(b) The matrices N_R *and* N_L *also have rank-2, with* $\mathsf{N}_R = [\boldsymbol{\rho}''_{11}, \boldsymbol{\rho}''_{12}, \boldsymbol{\rho}''_{13}]$ *and* $\mathsf{N}_L = [\boldsymbol{\rho}'''_{11}, \boldsymbol{\rho}'''_{12}, \boldsymbol{\rho}'''_{13}]$*. The left nullspace of* N_R *is the epipole* $\mathbf{e}''_1$ *and the left nullspace of* N_L *is the epipole* $\mathbf{e}'''_1$*.*

(c) Any line $\boldsymbol{\lambda}_{e''_1} \neq \boldsymbol{\rho}''_{1i}$ *through the epipole* $\mathbf{e}''_1$ *is mapped by* l_i *to the same point in image* Ψ_3 *– the epipole* $\mathbf{e}'''_1$*.*

Proof:
The property (a) follows directly from the nullspaces of rank-2 line correlations (see table 6.2 on page 58), which are the maps of the inducing 3D line into the images. Since the correlation l_i of the trifocal tensor are induced by the ith principal ray $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1, the map of this ray, which are the lines $\boldsymbol{\rho}''_{1i}$ and $\boldsymbol{\rho}'''_{1i}$ in image Ψ_2 and image Ψ_2 respectively, represent the nullspaces. The relation to the columns of the fundamental matrices is due to proposition 6.4 on page 47.

The property (b) is derived by observing that in image Ψ_2 all epipolar lines with respect to image Ψ_1 have to pass through the respective epipole (i.e. $\mathbf{e}''_1$). Therefore they form a pencil of epipolar lines with carrier $\mathbf{e}''_1$ and the matrix N_R, whose columns are made up of such concurrent lines can at most have rank-2. The same reasoning can be applied for the matrix N_L. We could also argue that the principal rays of image Ψ_1 are concurrent in space, consequently their maps must be concurrent in the images.

The correctness of property (c) can been seen simply by the fact, that the intersection point of the projection plane for any line $\boldsymbol{\lambda}_{e''_1} \neq \boldsymbol{\rho}''_{1i}$ through the epipole $\mathbf{e}''_1$ with $\boldsymbol{\mathcal{R}}'_i$ is the projection center of the first image, which maps to the epipole $\mathbf{e}'''_1$ in the third image. Algebraically we can prove it in the following way.

Let $\boldsymbol{\lambda}_{e''_1}$ be represented as $\boldsymbol{\lambda}_{e''_1} \sim \mathsf{S}(\boldsymbol{e}''_1)\boldsymbol{x}''$ with $\mathbf{x}'' \neq \mathbf{e}''_1$ and $\mathbf{x}'' \notin \boldsymbol{\rho}''_{1i}$. Then using the equation (7.10) on page 66 for l_i we get $\mathsf{l}_i\boldsymbol{\lambda}_{e''_1} \sim \mathbf{e}'''_1\, \boldsymbol{i}_i{}^\top \mathsf{A}^\top \mathsf{S}(\boldsymbol{e}''_1)\boldsymbol{x}'' = \mathbf{e}'''_1 s$. This shows that the result will always be $\sim \mathbf{e}'''_1$ as long as the scalar s is non-zero. This scalar s can be written as the determinant $|\mathsf{M}_1\boldsymbol{i}_i,\ \boldsymbol{Z}_1 - \boldsymbol{Z}_2,\ \mathsf{M}_2{}^{-1}\boldsymbol{x}''|$, where $\mathsf{M}_1\boldsymbol{i}_i$ is the direction of the ith principal ray $\boldsymbol{\mathcal{R}}'_i$ of the image Ψ_1. After some simple but tedious manipulations it can be seen that this determinant is non-zero as long as $\boldsymbol{Z}_2 \notin \boldsymbol{\mathcal{R}}'_i$ (i.e. $\operatorname{rank}(\mathsf{l}_i) = 2$), and $\mathbf{x}'' \neq \mathbf{e}''_1$ and $\mathbf{x}'' \notin \boldsymbol{\rho}''_{1i}$ (i.e. our assumptions on the point $\mathbf{x}''$).

□

The properties (a) and (b) in proposition 7.4 have a simple geometric interpretation. The column vectors of the correlation slices $\mathsf{l}_1 = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3\end{bmatrix}$, $\mathsf{l}_2 = \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3\end{bmatrix}$ and $\mathsf{l}_3 = \begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3\end{bmatrix}$ can be interpreted as points in image Ψ_3. Due to the singularity of each l_i matrix the respective three points have to lie on a line, which is the nullspace $\boldsymbol{\rho}'''_i$. And all three nullspaces $\boldsymbol{\rho}'''_1$, $\boldsymbol{\rho}'''_2$, $\boldsymbol{\rho}'''_3$ must be concurrent through the epipole $\mathbf{e}'''_1$; see figure 7.4 on the following page. The same geometric interpretation in image Ψ_2 holds for the $\mathsf{l}_i^\top$ matrices, which then involves the row vectors of l_i and the right nullspaces $\boldsymbol{\rho}''_i$.

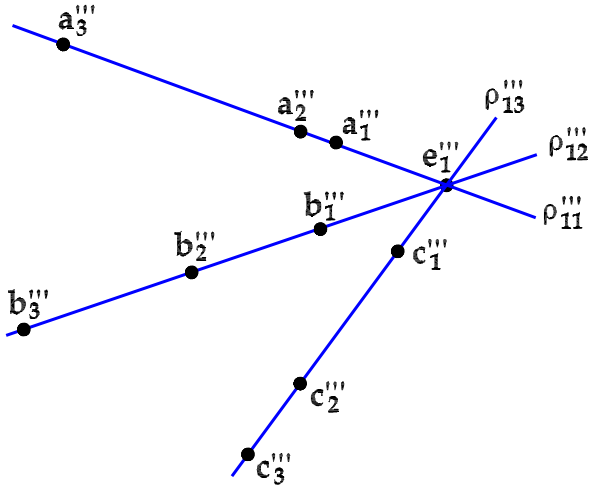

Figure 7.4: *The column vectors of the correlation slices* l_1*,* l_2 *and* l_3 *interpreted as points in image* Ψ_3*.*

7.2.2 The homography slices

Next we will fix the j-index in the trifocal tensor (7.7).

The J homography slices Fixing the j-index returns a two-fold indexed quantity $(J_j)^k_{\cdot i}$

$$(J_j)^k_{\cdot i} = T^{k(j)}_{\cdot\ \cdot\ i} = e_1''^{(j)} B^k_{\cdot i} - e_1'''^k A^{(j)}_{\cdot\ i}.$$

Again we can interpret this as a 3×3 matrix J_j, where the rows are created by the first index (k) and the columns by the second index (i). Therefore $A^{(j)}_{\cdot\ i}$ corresponds to $\boldsymbol{i}_j{}^\top\mathsf{A}$ and $B^k_{\cdot i}$ to B. And so we get J_j in matrix notation as

$$\mathsf{J}_j = (\mathbf{e}_1''^\top\ \boldsymbol{i}_j)\ \mathsf{B} - \mathbf{e}_1'''\ \boldsymbol{i}_j{}^\top\mathsf{A}. \tag{7.13}$$

We can further express J_j in terms of the projection matrices $\mathsf{P}_h \sim \mathsf{M}_h\,[I_3, -\boldsymbol{Z}_h]$ with regular M_h using the relations (7.4) and (7.2) as

$$\mathsf{J}_j = \mathsf{M}_3\Big(\underbrace{(\boldsymbol{Z}_1 - \boldsymbol{Z}_2)^\top\mathsf{M}_2{}^\top\boldsymbol{i}_j}_{d}\,I + (\boldsymbol{Z}_3 - \boldsymbol{Z}_1)\,\underbrace{\boldsymbol{i}_j{}^\top\mathsf{M}_2}_{\boldsymbol{\Pi}''_{j,H}{}^\top}\Big)\mathsf{M}_1{}^{-1}, \tag{7.14}$$

where $d = \boldsymbol{Z}_1{}^\top\boldsymbol{\Pi}''_{j,H} + \Pi''_{j,O}$ and $\boldsymbol{\Pi}''_j{}^\top = (\boldsymbol{\Pi}''_{j,H}{}^\top, \Pi''_{2j,O})$ represents the jth principal plane $\boldsymbol{\Pi}''_i$ of image Ψ_2; see relation 5.13 on page 39. The comparison of the equation (7.14) on page 70 with proposition 6.7 on page 53 reveals that J_j describes a *homography* of the points in image Ψ_1 to the points in image Ψ_3 via the jth principal plane $\boldsymbol{\Pi}''_i$ of image Ψ_2; see figure 7.5 on the next page. For this reason we will refer to these slices as **homography slices**[2] and summarize these findings in the following proposition.

[2] In [Faugeras and Luong 2001] the respective slices are called *intrinsic planar morphisms* H^v_{ji}. Due to their choice of indices these matrices correspond to the matrices $\mathsf{J}_j{}^\top$ considered in this thesis.

Proposition 7.5 (The homography slices J_j of the trifocal tensor.) *Let the trifocal tensor be given as $T_{\cdot\cdot i}^{kj} \sim e_1''^{j} B_{\cdot i}^{k} - e_1'''^{k} A_{\cdot i}^{j}$. Then the homography slices J_j, which are obtained by fixing the j-index, are represented as $\mathsf{J}_j = (\mathbf{e}_1''^{\top} \boldsymbol{i}_j)\, \mathsf{B} - \mathbf{e}_1''' \boldsymbol{i}_j{}^{\top} \mathsf{A}$. The matrix J_j represents a collineation of the points in image Ψ_1 to the points in image Ψ_3 via the jth principal plane $\mathbf{\Pi}_i''$ of image Ψ_2. The matrix $\mathsf{J}_j{}^{\top}$ maps the lines in image Ψ_3 to the lines in image Ψ_1 via the very same principal plane.*

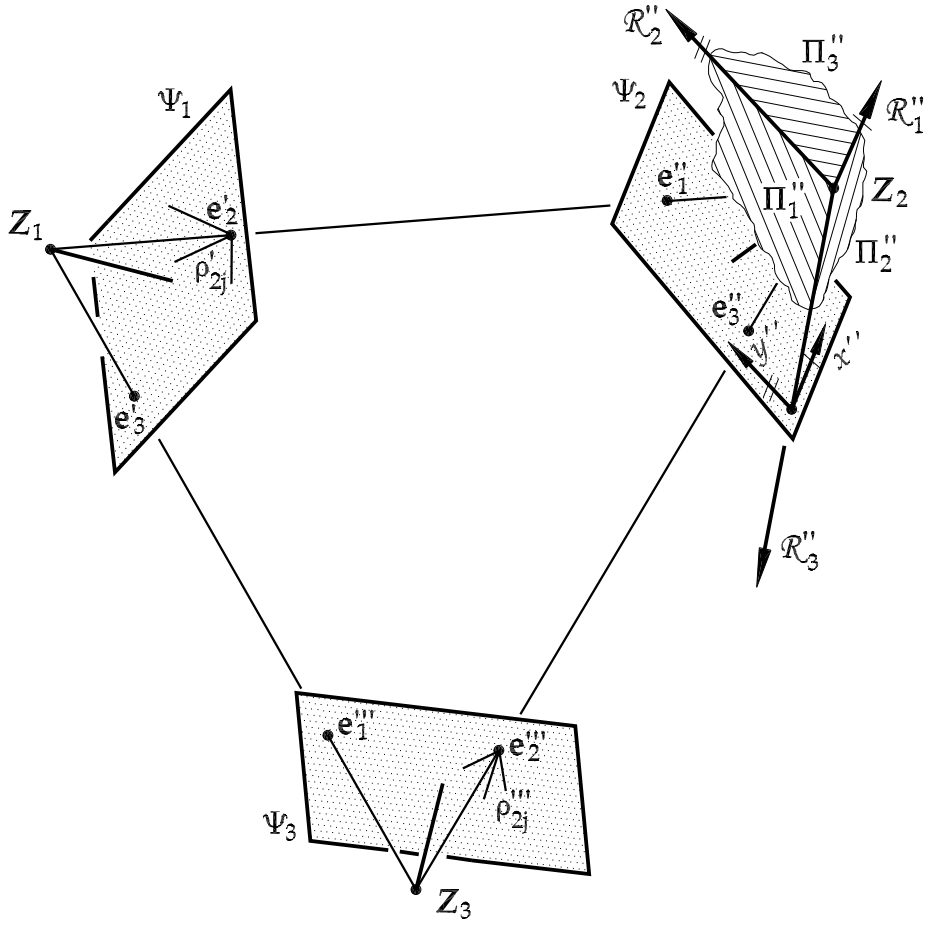

Figure 7.5: *The J_j matrices represent plane homographies from image Ψ_1 to image Ψ_3 induced by the principal planes $\mathbf{\Pi}_j''$ of image Ψ_2. These principal planes intersect in the principal rays $\boldsymbol{\mathcal{R}}_j''$ and their maps in the images Ψ_1 and Ψ_3 are the lines $\boldsymbol{\rho}_{2j}'$ and $\boldsymbol{\rho}_{2j}'''$ respectively.*

Proposition 7.6 (Linear combinations of the homography slices J_j.) *The matrix $\mathsf{J}_\diamond(\boldsymbol{\lambda}'')$, which is the linear combination of the homography slices J_j given by*

$$\mathsf{J}_\diamond(\boldsymbol{\lambda}'') = a''\mathsf{J}_1 + b''\mathsf{J}_2 + c''\mathsf{J}_3, \tag{7.15}$$

or

$$\mathsf{J}_\diamond(\boldsymbol{\lambda}'') = \mathsf{M}_3\Big((\boldsymbol{Z}_1 - \boldsymbol{Z}_2)^{\top}\mathsf{M}_2{}^{\top}\boldsymbol{\lambda}''\boldsymbol{I} + (\boldsymbol{Z}_3 - \boldsymbol{Z}_1)\boldsymbol{\lambda}''^{\top}\mathsf{M}_2\Big)\mathsf{M}_1{}^{-1},$$

describes the collineation of the points in image Ψ_1 to the points in image Ψ_3 via the projection plane of the line $\boldsymbol{\lambda}'' \sim (a'', b'', c'')^{\top}$ in image Ψ_2.

Proof:
The comparison of the second expression for $\mathsf{J}_\diamond(\boldsymbol{\lambda}'')$ with relation (7.14) shows that $\mathsf{J}_\diamond(\boldsymbol{\lambda}'')$ corresponds to the homography matrix with respect to a 3D plane passing through $\mathbf{Z}_2$ and whose normal vector is given by $\mathsf{M}_2{}^\top\boldsymbol{\lambda}''$; i.e. the projection plane of the line $\boldsymbol{\lambda}''$ in the image Ψ_2.

□

Adapting from proposition 6.8 on page 54 we can give the rank of J_j depending on the arrangement of the three images Ψ_1, Ψ_2 and Ψ_3.

Proposition 7.7 (The ranks of the homography slices J_j.) *In general the rank of J_j is 3 and the epipoles $\mathbf{e}'_3$ (in image Ψ_1) and $\mathbf{e}'''_1$ (in image Ψ_3) are related by*

$$\mathbf{e}'''_1 \sim \mathsf{J}_j\mathbf{e}'_3. \tag{7.16}$$

The rank of J_j is 2, if the jth principal plane $\boldsymbol{\Pi}''_j$ of image Ψ_2 contains the projection center $\mathbf{Z}_3$ of image Ψ_3; then it holds: $e''^{(j)}_3 = 0$.
The rank of J_j is 1, if the jth principal plane $\boldsymbol{\Pi}''_j$ of image Ψ_2 contains the projection center $\mathbf{Z}_1$ of image Ψ_1; then it holds: $e''^{(j)}_1 = 0$.
The rank of J_j is 0, if the jth principal plane $\boldsymbol{\Pi}''_j$ of image Ψ_2 contains the coinciding projection centers $\mathbf{Z}_1 = \mathbf{Z}_3$; then it holds: $\mathbf{e}'_3 = \mathbf{e}'''_1 = \mathbf{0}$.
Depending on its rank, *J_j has the ranges and nullspaces shown in table 7.2.*

$\mathrm{rank}(\mathsf{J}_j)$	$\mathcal{R}(\mathsf{J}_j)$	$\mathcal{N}(\mathsf{J}_j)$
3	$\mathbb{P}^2$	$\{\}$
2	$\boldsymbol{\sigma}''' \sim \mathsf{B}^{-\top}\mathsf{A}^\top\boldsymbol{i}_j$	$\mathbf{e}'_3$
1	$\mathbf{e}'''_1$	$\boldsymbol{\sigma}' \sim \mathsf{A}^\top\boldsymbol{i}_j$
0	$\{\}$	$\mathbb{P}^2$

Table 7.2: *Range $\mathcal{R}(\mathsf{J}_j)$ and nullspace $\mathcal{N}(\mathsf{J}_j)$ of the homography slice J_j depending on its rank. The lines $\boldsymbol{\sigma}'$ and $\boldsymbol{\sigma}'''$ arise as the intersection of the jth principal plane $\boldsymbol{\Pi}''_j$ of image Ψ_2 with the other two images and are given using the defining matrices $\mathsf{A} = \mathsf{M}_2\mathsf{M}_1{}^{-1}$ and $\mathsf{B} = \mathsf{M}_3\mathsf{M}_1{}^{-1}$ of the trifocal tensor.*

Proof:
All these properties are obtained directly from proposition 6.8 except for the additional properties of the vanishing epipolar coordinates for the singular cases. These properties, however, are easily verified by observing that for singular J_j matrices it holds:

$$\begin{aligned}
\mathrm{rank}(\mathsf{J}_j) = 2 \quad &\Leftrightarrow \quad \mathbf{Z}_3 \in \boldsymbol{\Pi}''_j \Rightarrow 0 = \boldsymbol{\Pi}''_{j,H}{}^\top\mathbf{Z}_3 + \Pi_{2j,O} = \boldsymbol{i}_j{}^\top\mathsf{M}_2(\mathbf{Z}_3 - \mathbf{Z}_2) = \boldsymbol{i}_j{}^\top\mathbf{e}''_3 \\
\mathrm{rank}(\mathsf{J}_j) = 1 \quad &\Leftrightarrow \quad \mathbf{Z}_1 \in \boldsymbol{\Pi}''_j \Rightarrow 0 = \boldsymbol{\Pi}''_{j,H}{}^\top\mathbf{Z}_1 + \Pi_{2j,O} = \boldsymbol{i}_j{}^\top\mathsf{M}_2(\mathbf{Z}_3 - \mathbf{Z}_1) = \boldsymbol{i}_j{}^\top\mathbf{e}''_1
\end{aligned}$$

The form of $\mathbf{s}'$ and $\mathbf{s}''$ follows from table 6.1 on page 54 with relation (5.13) for the normal vector of the principal plane $\boldsymbol{\Pi}''_j$ and the defining matrices of the trifocal tensor in (7.4).

□

From equation (7.16) we see that $\mathbf{e}'_3$ is a solution of the general eigenvalue problem $(\mathsf{J}_p - \kappa\mathsf{J}_q)\boldsymbol{x} = \mathbf{0}$. The following proposition goes more into detail.

Proposition 7.8 (The general eigenvalue problem $(\mathsf{J}_p - \kappa\mathsf{J}_q)\mathbf{x}' = \mathbf{0}$.) *The general eigenvalue problem $(\mathsf{J}_p - \kappa\mathsf{J}_q)\mathbf{x}' = \mathbf{0}$, with regular J_q, $\mathbf{x}' \in \Psi_1$ and $p \neq q$, has the following general eigenvalues:*

- *A double general eigenvalue $\kappa_1 = \kappa_2 = e_1''^{(p)}/e_1''^{(q)}$ with the corresponding two dimensional general eigenspace $\boldsymbol{\rho}_{2r}'^{\top}\mathbf{x}' = 0$, where $\boldsymbol{\rho}_{2r}' \sim \mathsf{M}_1^{-\top}\mathsf{M}_2^{\top}\mathsf{S}(\mathbf{e}_1'')\boldsymbol{i}_r$ is the map of the r*th *principal ray $\boldsymbol{\mathcal{R}}_r''$ of image Ψ_2 into image Ψ_1 ($r \neq p$ and $r \neq q$); see figure 7.5 on page 71.*
- *A single general eigenvalue $\kappa_3 = e_3''^{(p)}/e_3''^{(q)}$ with the corresponding general eigenvector $\mathbf{e}_3'$.*

This general eigenvalue problem may have an eigenvalue with multiplicity 3 ($\kappa_1 = \kappa_2 = \kappa_3 = e_1''^{(p)}/e_1''^{(q)}$). If the respective rank defect is 3 then it holds $\mathbf{Z}_1 = \mathbf{Z}_3$, otherwise $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_3$ and $\boldsymbol{\mathcal{R}}_r''$ are coplanar and $\mathbf{e}_3' \in \boldsymbol{\rho}_{2r}'$. If the latter case is true for two different choices of $\{p, q\}$, then $\mathbf{Z}_1$, $\mathbf{Z}_2$ and $\mathbf{Z}_3$ are collinear.

If J_p is singular with rank-2, then $\kappa_3 = 0$. If J_p is singular with rank 1, then $\kappa_1 = \kappa_2 = 0$; if also one of the previous condition holds then $\kappa_1 = \kappa_2 = \kappa_3 = 0$.

Proof:
First we will show that $e_1''^{(p)}/e_1''^{(q)}$ is a double root. Using the expression (7.13) for J_j we get

$$\mathsf{J}_p - \kappa\mathsf{J}_q = \mathbf{e}_1''^{\top}(\boldsymbol{i}_p - \kappa\boldsymbol{i}_q)\mathsf{B} - \mathbf{e}_3'(\boldsymbol{i}_p - \kappa\boldsymbol{i}_q)^{\top}\mathsf{A}.$$

From this relation it is easy to see that for $\kappa_1 = e_1''^{(p)}/e_1''^{(q)}$ the first addend vanishes and only the second one – a dyad – which has rank 1 remains. Since the multiplicity of an eigenvalue is greater or equal to the rank defect (see inequality (B.4) on page 177), which is 2, κ_1 must be a at least a double root. The corresponding two dimensional eigenspace can be found by expressing $(\mathsf{J}_p - \kappa_1\mathsf{J}_q)\mathbf{x}' = \mathbf{0}$ in the defining quantities of the trifocal tensor using relations (7.2) - (7.4), which gives

$$\mathsf{M}_3(\boldsymbol{Z}_1 - \boldsymbol{Z}_3)\boldsymbol{i}_r^{\top}\mathsf{S}(\mathbf{e}_1'')\mathsf{M}_2\mathsf{M}_1^{-1}\mathbf{x}' = \mathsf{M}_3(\boldsymbol{Z}_1 - \boldsymbol{Z}_3)\boldsymbol{\rho}_{2r}'^{\top}\mathbf{x}' = \mathbf{0}. \tag{7.17}$$

So, as long as $\mathbf{Z}_1 \neq \mathbf{Z}_3$, the line $\boldsymbol{\rho}_{2r}' \sim \mathsf{M}_1^{-\top}\mathsf{M}_2^{\top}\mathsf{S}(\mathbf{e}_1'')\boldsymbol{i}_r$ in image Ψ_1 represents the general eigenspace. To see that $\boldsymbol{\rho}_{2r}'$ is the map of the rth principal ray $\boldsymbol{\mathcal{R}}_r''$ of image Ψ_2 into image Ψ_1, we can argue in the following way. The expression $\mathsf{M}_2^{\top}\mathsf{S}(\mathbf{e}_1'')\boldsymbol{i}_r$ represents the normal vector $\boldsymbol{n}_\Sigma$ of the second projection plane $\boldsymbol{\Sigma}$ of a line joining $\mathbf{e}_1''$ and $\boldsymbol{i}_r$; see relations (3.24) and (5.21). The rth principal ray $\boldsymbol{\mathcal{R}}_r''$ also lies in this plane as it is defined by the second projection center and its direction (in the system of the second camera) is given by $\boldsymbol{i}_r$. Further, since $\mathbf{e}_1''$ lies in this plane also the first projection center has to. Therefore $\boldsymbol{\Sigma}$ is also a projection plane for the first image and the respective image line is given by $\mathsf{M}_1^{-\top}\boldsymbol{n}_\Sigma$.

Now for the single eigenvalue. From equation (7.16) we see that the epipole $\mathbf{e}_3'$ is also a general eigenvector, which in general should not lie on $\boldsymbol{\rho}_{2r}'$ because the latter is independent of the image Ψ_3. Inserting $\mathbf{x}' \sim \mathbf{e}_3' = \mathsf{M}_1(\boldsymbol{Z}_3 - \boldsymbol{Z}_1)$ in $(\mathsf{J}_p - \kappa\mathsf{J}_q)\mathbf{x}' = \mathbf{0}$ gives

$$\mathsf{M}_3(\boldsymbol{Z}_3 - \boldsymbol{Z}_1)(\boldsymbol{i}_p - \kappa\boldsymbol{i}_q)^{\top}\mathsf{M}_2(\boldsymbol{Z}_3 - \boldsymbol{Z}_2) = \mathbf{0}.$$

And so, as long as $\mathbf{Z}_1 \neq \mathbf{Z}_3$, $\mathbf{e}_3'$ is the general eigenvector to the single eigenvalue $\kappa_3 = e_3''^{(p)}/e_3''^{(q)}$.

Since the denominator in all general eigenvalues must be non-zero, the matrix J_q is required to have maximum rank; see the relations for the vanishing epipole coordinates in proposition 7.7 on the preceding page.

In case of one eigenvalue having multiplicity 3 the rank defect in $\mathsf{J}_p - \kappa\mathsf{J}_q$ can become 3; then $\mathsf{J}_p - \kappa\mathsf{J}_q = \mathsf{0}$ and therefore $\mathbf{Z}_1 = \mathbf{Z}_3$. In general the rank defect will be 2 and then $\mathbf{e}'_3$ is no longer a unique eigenvector but lies in $\boldsymbol{\rho}'_{2r}$. If this later case occurs for two different choices of $\{p, q\}$ then $\mathbf{e}'_3$ has to lie on two different $\boldsymbol{\rho}'_{2r}$ which intersect in $\mathbf{e}'_2$ and thus $\mathbf{e}'_2 = \mathbf{e}'_3$ and the three projection centers are collinear.

The properties for singular J_p follow from proposition 7.7.

□

The fact that the rth principal ray $\boldsymbol{\mathcal{R}}''_r$ of image Ψ_2 is responsible for one eigenspace can geometrically be seen by rewriting the general eigenvalue problem $(\mathsf{J}_p - \kappa\mathsf{J}_q)\mathbf{x}' = \mathbf{0}$ as a standard one $({\mathsf{J}_q}^{-1}\mathsf{J}_p - \kappa\mathsf{I})\mathbf{x}' = \mathbf{0}$, which is possible since J_q is required to have maximum rank. Geometrically this means, we map a point $\mathbf{x}'$ from image Ψ_1 via the homography plane of J_p into image Ψ_3 and then project this point back from image Ψ_3 into image Ψ_1 via the homography plane of J_q. Obviously we only get the point $\mathbf{x}'$ from where we started, if $\mathbf{x}'$ corresponds to a point on the intersection line of the two homography planes. Now, since these planes are the pth and qth principal plane of image Ψ_2, they intersect in the rth principal ray of image Ψ_2.

The K homography slices By fixing the k-index we get a two-fold indexed quantity $(K_k)^j_{\cdot i}$

$$(K_k)^j_{\cdot i} = T^{(k)j}_{\cdot\ \cdot i} = e''^j_1 B^{(k)}_{\cdot\ i} - e'''^{(k)}_1 A^j_{\cdot i}.$$

Again we can interpret this as a 3×3 matrix K_k, where the rows are created by the first index (j) and the columns by the second index (i). Therefore $A^j_{\cdot i}$ corresponds to A and $B^{(k)}_{\cdot\ i}$ to $\boldsymbol{i}_k{}^\top\mathsf{B}$. And so we get K_k in matrix notation as

$$\mathsf{K}_k = \mathbf{e}''_1\boldsymbol{i}_k^\top\mathsf{B} - (\mathbf{e}'''^\top_1\boldsymbol{i}_k)\ \mathsf{A}. \tag{7.18}$$

We can further express K_k in terms of the projection matrices $\mathsf{P}_h \sim \mathsf{M}_h\,[\mathsf{I}_3, -\mathbf{Z}_h]$ with regular M_h using the relations (7.4) and (7.2) as

$$\mathsf{K}_k = -\mathsf{M}_2\Big(\underbrace{(\mathbf{Z}_1 - \mathbf{Z}_3)^\top\mathsf{M}_3{}^\top\boldsymbol{i}_k}_{d}\,\mathsf{I} + (\mathbf{Z}_2 - \mathbf{Z}_1)\,\underbrace{\boldsymbol{i}_k{}^\top\mathsf{M}_3}_{\boldsymbol{\Pi}'''^\top_{i,H}}\Big)\mathsf{M}_1{}^{-1}, \tag{7.19}$$

where $d = \mathbf{Z}_1{}^\top\boldsymbol{\Pi}'''_{k,H} + \Pi'''_{k,O}$ and $\boldsymbol{\Pi}'''^\top_k = (\boldsymbol{\Pi}'''_{k,H}{}^\top, \Pi'''_{k,O})$ represents the kth principal plane $\boldsymbol{\Pi}'''_k$ of image Ψ_3; see relation 5.13 on page 39. The comparison of the equation (7.19) on page 74 with proposition 6.7 on page 53 reveals that K_k describes also a *homography* just as the J slices, but of the points in image Ψ_1 to the points in image Ψ_2 via the kth principal plane $\boldsymbol{\Pi}'''_k$ of image Ψ_3; see figure 7.6 on the facing page. So we see that the K_k and J_j matrices represent mappings of the same kind. Therefore we can adopt the properties from the J_j matrices, by interchanging the images Ψ_2 and Ψ_3, and give the following properties.

Proposition 7.9 (The homography slices K_k of the trifocal tensor.) *Let the trifocal tensor be given as $T^{kj}_{\cdot\cdot i} \sim e''^j_1 B^k_{\cdot i} - e'''^k_1 A^j_{\cdot i}$. Then the homography slices K_k, which are obtained by fixing the k-index, are represented as $\mathsf{K}_k = \mathbf{e}''_1\boldsymbol{i}_k{}^\top\mathsf{B} - (\mathbf{e}'''^\top_1\boldsymbol{i}_k)\mathsf{A}$. The matrix K_k represents a collineation of the points in image Ψ_1 to the points in image Ψ_2 via the kth principal plane $\boldsymbol{\Pi}'''_k$ of image Ψ_3. The matrix $\mathsf{K}_k{}^\top$ maps the lines in image Ψ_2 to the lines in image Ψ_1 via the very same principal plane.*

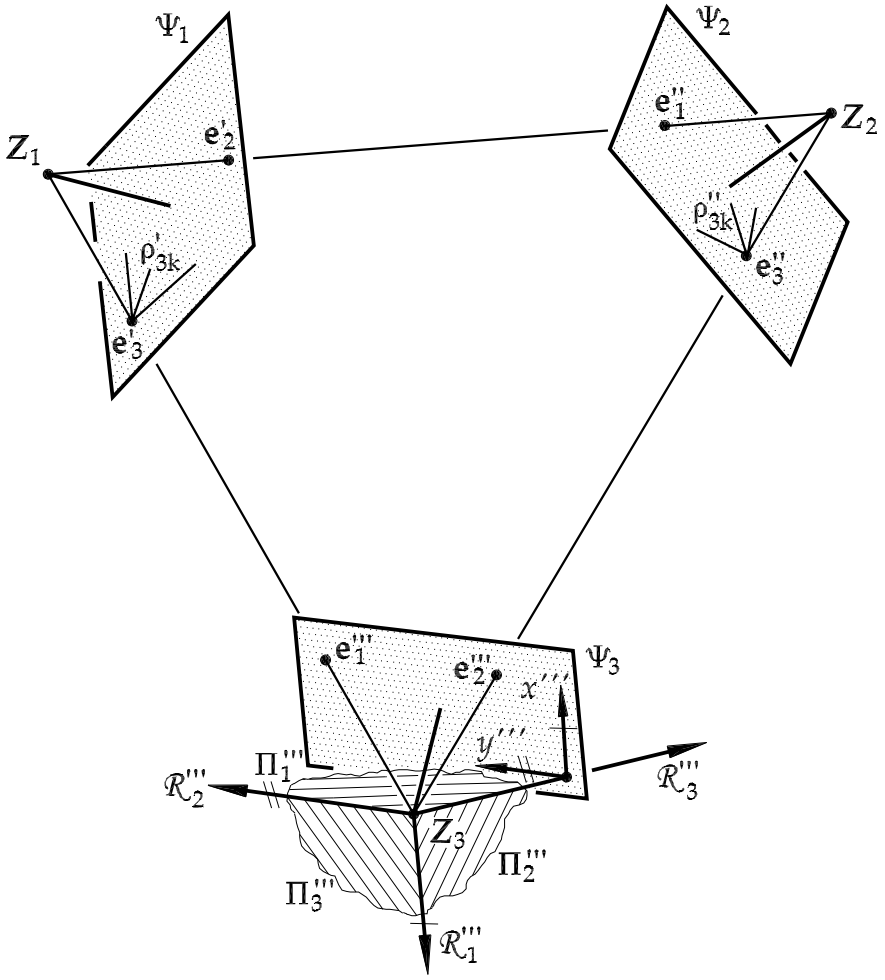

Figure 7.6: *The K_k matrices represent plane homographies from image Ψ_1 to image Ψ_2 induced by the principal planes $\mathbf{\Pi}'''_k$ of image Ψ_3. These principal planes intersect in the principal rays $\boldsymbol{\mathcal{R}}'''_k$ and their maps in the images Ψ_1 and Ψ_2 are the lines $\boldsymbol{\rho}'_{3k}$ and $\boldsymbol{\rho}''_{3k}$ respectively.*

Proposition 7.10 (Linear combinations of the homography slices K_k.) *The matrix $\mathsf{K}_\diamond(\boldsymbol{\lambda}''')$, which is the linear combination of the homography slices K_k given by*

$$\mathsf{K}_\diamond(\boldsymbol{\lambda}''') = a'''\mathsf{K}_1 + b'''\mathsf{K}_2 + c'''\mathsf{K}_3, \tag{7.20}$$

or

$$\mathsf{K}_\diamond(\boldsymbol{\lambda}''') = -\mathsf{M}_2\Big((\boldsymbol{Z}_1 - \boldsymbol{Z}_3)^\top \mathsf{M}_3{}^\top \boldsymbol{\lambda}''' I + (\boldsymbol{Z}_2 - \boldsymbol{Z}_1)\boldsymbol{\lambda}'''^\top \mathsf{M}_3\Big)\mathsf{M}_1{}^{-1},$$

describes the collineation of the points in image Ψ_1 to the points in image Ψ_2 via the projection plane of the line $\boldsymbol{\lambda}''' \sim (a''', b''', c''')^\top$ in image Ψ_3.

Proposition 7.11 (The ranks of the homography slices K_k.) *In general the rank of K_k is 3 and the epipoles $\mathbf{e}'_2$ (in image Ψ_1) and $\mathbf{e}''_1$ (in image Ψ_2) are related by*

$$\mathbf{e}''_1 \sim \mathsf{K}_k \mathbf{e}'_2. \tag{7.21}$$

The rank of K_k *is 2, if the* kth *principal plane* $\mathbf{\Pi}_k'''$ *of image* Ψ_3 *contains the projection center* $\mathbf{Z}_2$ *of image* Ψ_2*; then it holds:* $e_2'''^{(k)} = 0$.
The rank of K_k *is 1, if the* kth *principal plane* $\mathbf{\Pi}_k'''$ *of image* Ψ_3 *contains the projection center* $\mathbf{Z}_1$ *of image* Ψ_1*; then it holds:* $e_1'''^{(k)} = 0$.
The rank of K_k *is 0, if the* kth *principal plane* $\mathbf{\Pi}_k'''$ *of image* Ψ_3 *contains the coinciding projection centers* $\mathbf{Z}_1 = \mathbf{Z}_2$*; then it holds:* $\mathbf{e}_2' = \mathbf{e}_1'' = \mathbf{0}$.
Depending on its rank, K_k *has the ranges and nullspaces shown in table 7.3.*

$\operatorname{rank}(\mathsf{K}_k)$	$\mathcal{R}(\mathsf{K}_k)$	$\mathcal{N}(\mathsf{K}_k)$
3	$\mathbb{P}^2$	$\{\}$
2	$\boldsymbol{\sigma}'' \sim \mathsf{A}^{-\top}\mathsf{B}^\top \boldsymbol{i}_j$	$\mathbf{e}_2'$
1	$\mathbf{e}_1''$	$\boldsymbol{\sigma}' \sim \mathsf{B}^\top \boldsymbol{i}_j$
0	$\{\}$	$\mathbb{P}^2$

Table 7.3: *Range* $\mathcal{R}(\mathsf{K}_k)$ *and nullspace* $\mathcal{N}(\mathsf{K}_k)$ *of the homography slice* K_k *depending on its rank. The lines* $\boldsymbol{\sigma}'$ *and* $\boldsymbol{\sigma}''$ *arise as the intersection of the* kth *principal plane* $\mathbf{\Pi}_k'''$ *of image* Ψ_3 *with the other two images and are given using the defining matrices* $\mathsf{A} = \mathsf{M}_2\mathsf{M}_1{}^{-1}$ *and* $\mathsf{B} = \mathsf{M}_3\mathsf{M}_1{}^{-1}$ *of the trifocal tensor.*

Proposition 7.12 (The general eigenvalue problem $(\mathsf{K}_p - \kappa\mathsf{K}_q)\mathbf{x}' = \mathbf{0}$**.)** *The general eigenvalue problem* $(\mathsf{K}_p - \kappa\mathsf{K}_q)\mathbf{x}' = \mathbf{0}$*, with regular* K_q*,* $\mathbf{x}' \in \Psi_1$ *and* $p \neq q$*, has the following general eigenvalues:*

- *A double general eigenvalue* $\kappa_1 = \kappa_2 = e_1''^{(p)}/e_1''^{(q)}$ *with the corresponding two dimensional general eigenspace* $\boldsymbol{\rho}_{3r}'^{\top}\mathbf{x}' = 0$*, where* $\boldsymbol{\rho}_{3r}' \sim \mathsf{M}_1{}^{-\top}\mathsf{M}_3{}^\top\mathsf{S}(\mathbf{e}_1''')\boldsymbol{i}_r$ *is the map of the* rth *principal ray* $\boldsymbol{\mathcal{R}}_r'''$ *of image* Ψ_3 *into image* Ψ_1 *(*$r \neq p$ *and* $r \neq q$*); see figure 7.6 on the page before.*
- *A single general eigenvalue* $\kappa_3 = e_2'''^{(p)}/e_2'''^{(q)}$ *with the corresponding general eigenvector* $\mathbf{e}_2'$.

This general eigenvalue problem may have an eigenvalue with multiplicity 3 ($\kappa_1 = \kappa_2 = \kappa_3 = e_1'''^{(p)}/e_1'''^{(q)}$*). If the respective rank defect is 3 then it holds* $\mathbf{Z}_1 = \mathbf{Z}_3$*, otherwise* $\mathbf{Z}_1$*,* $\mathbf{Z}_2$*,* $\mathbf{Z}_3$ *and* $\boldsymbol{\mathcal{R}}_r'''$ *are coplanar and* $\mathbf{e}_2' \in \boldsymbol{\rho}_{3r}'$*. If the latter case is true for two different choices of* $\{p, q\}$*, then* $\mathbf{Z}_1$*,* $\mathbf{Z}_2$ *and* $\mathbf{Z}_3$ *are collinear.*

If K_p *is singular with rank-2, then* $\kappa_3 = 0$*. If* K_p *is singular with rank 1, then* $\kappa_1 = \kappa_2 = 0$*; if also one of the previous condition holds then* $\kappa_1 = \kappa_2 = \kappa_3 = 0$.

7.2.3 Rank combinations

In the two previous sections we looked at the properties of the tensor slices. These properties are directly linked to the ranks of these slices. For the correlation slices rank defects occur if the projection centers $\mathbf{Z}_2$ or $\mathbf{Z}_3$ lie on the principal rays of the first image, whereas for the homography slices they occur if the projection centers $\mathbf{Z}_1$, $\mathbf{Z}_2$ resp. $\mathbf{Z}_3$ lie in the principal planes of the third resp. second image. Therefore we should know what possible rank defects can occur depending on the arrangement of the projection centers with respect to the principal rays

and planes. Note that the orientation of the principal rays and planes of an image is defined by the 3D image coordinate system (see figure 5.3 on page 38), which itself depends on the image measurement system. Therefore we can control the rank defects of the tensor slices by adapting the image coordinate systems.

Let us first consider the usual case of distinct projection centers $\mathbf{Z}_1 \neq \mathbf{Z}_2 \neq \mathbf{Z}_3$. Because the projection centers $\mathbf{Z}_2$ and $\mathbf{Z}_3$ can not lie on all three principal rays of the first image at the same time, at least one rank-2 correlation slice is available, with the others having at least rank-1. Now for the homography slices: If $\mathbf{Z}_1$ lies on the principal ray $\boldsymbol{\mathcal{R}}_1''$ of image Ψ_2 then the J_2 and J_3 have rank 1, as these homographies are induced by the principal planes $\boldsymbol{\Pi}_2''$ and $\boldsymbol{\Pi}_3''$, all of which contain $\mathbf{Z}_1$ in this case. Only J_1 would be regular, however, if $\mathbf{Z}_3$ lies on one of the other two principal rays of image Ψ_2 or in the principal plane $\boldsymbol{\Pi}_1''$ then also J_1 is singular with rank-2. Further one of the principal rays of image Ψ_3 could contain one of the other two projection centers, then also two K homographies would be singular. But at least one of the K homographies will be regular, as not all three principal planes or equivalently two principal rays of image Ψ_3 can contain the other two projection centers (provided the image coordinate systems have no or only small affinity, which for real images is generally fulfilled).

This shows that in the case of three distinct projection centers at least one rank-2 correlation slice is available, with the others having at least rank-1, and at least one regular homography of one kind and at least one rank-2 homography of the other kind. This gives the first row in table 7.4. The entries for the cases of incident projection center can be found by similar considerations.

PRC incidence	rank(I_i)	rank(J_j), rank(K_k)
$\mathbf{Z}_1 \neq \mathbf{Z}_2 \neq \mathbf{Z}_3$	all I_i: rank ≥ 1, at least one I_i with rank=2	all J_j: rank ≥ 1, at least one J_j with rank=2 all K_k: rank ≥ 1, at least one K_k with rank=3 or J_j and K_k distribution interchanged
$\mathbf{Z}_2 = \mathbf{Z}_3$	all I_i: rank ≥ 0, at least two I_i with rank=2	all J_j: $1 \leq$ rank ≤ 2 all K_k: $1 \leq$ rank ≤ 2 at least one J_j and one K_k with rank=2
$\mathbf{Z}_1 = \mathbf{Z}_2$	all I_i: rank=1	all J_j: rank=1 all K_k: rank ≥ 0, at least one K_k with rank=3
$\mathbf{Z}_1 = \mathbf{Z}_3$	all I_i: rank=1	all K_k: rank=1 all J_j: rank ≥ 0, at least one J_j with rank=3
$\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{Z}_3$	all I_i: rank=0	all J_j and all K_k: rank=0

Table 7.4: *Possible rank combinations for the tensor slices I_i, J_j and K_k depending on the incidence of the projection centers (PRC).*

7.3 The trilinearities

The derivation of the trifocal tensor in section 7.1 was actually motivated by finding two constraints for corresponding image lines $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$ and $\boldsymbol{\lambda}_3$ in three images. These two constraints can

be represented in two ways. The first representation is given by the relation (7.6)

$$\lambda'_i \sim \lambda''_j \lambda'''_k T^{kj}_{\cdot\cdot i}.$$

Using the correlation slices from section 7.2.1 this expression can simply be transferred from tensor notation to vector notation as

$$\boldsymbol{\lambda}' \sim \begin{pmatrix} \boldsymbol{\lambda}'''^{\top} \mathsf{l}_1 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_2 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_3 \boldsymbol{\lambda}'' \end{pmatrix}, \tag{7.22}$$

which are three equations minus one unknown scale factor (note the $\sim$ symbol). We will call this a **transfer relation**, because it can be used to predict the position of the line in image Ψ_1 given the corresponding lines in the other two images.

The second representation applies the identity check in $\mathbb{P}^2$ (see section 3.6) and removes the scale ambiguity by multiplying with the axiator $\mathsf{S}(\boldsymbol{\lambda}')$. In this way we get the following set of constraints; e.g. [Hartley 1994a]

$$\mathsf{S}(\boldsymbol{\lambda}') \begin{pmatrix} \boldsymbol{\lambda}'''^{\top} \mathsf{l}_1 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_2 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_3 \boldsymbol{\lambda}'' \end{pmatrix} = \mathbf{0}. \tag{7.23}$$

Since the identity check in $\mathbb{P}^2$ has 2 degrees of freedom, we get two independent constraints. We will refer to these constraints as **line-trilinearities**. They constitute a homogenous trilinear form; *tri-linear* since (7.23) is linear in the elements of the trifocal tensor and each coefficient of this form is a product of one coordinate from each of the three image lines; *homogenous* since the right side is the zero vector.

From the transfer relation (7.22) we can derive another single constraint, which is caused by the fact that for any point $\mathbf{x}'$ on $\boldsymbol{\lambda}'$ it must hold $\mathbf{x}'^{\top} \boldsymbol{\lambda}' = 0$, thus

$$\mathbf{x}'^{\top} \begin{pmatrix} \boldsymbol{\lambda}'''^{\top} \mathsf{l}_1 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_2 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_3 \boldsymbol{\lambda}'' \end{pmatrix} = \boldsymbol{\lambda}'''^{\top} \mathsf{l}_\diamond(\mathbf{x}') \boldsymbol{\lambda}'' = 0, \tag{7.24}$$

with $\mathsf{l}_\diamond(\mathbf{x}')$ from proposition 7.2. This trilinearity involves the corresponding image features of a point in image Ψ_1, and a line in image Ψ_2 and image Ψ_3. From this constraint we can derive even more constraints. First, from

$$\underbrace{\boldsymbol{\lambda}'''^{\top} \mathsf{l}_\diamond(\mathbf{x}')}_{\mathbf{x}''^{\top}} \boldsymbol{\lambda}'' = \boldsymbol{\lambda}'''^{\top} \underbrace{\mathsf{l}_\diamond(\mathbf{x}') \boldsymbol{\lambda}''}_{\mathbf{x}'''} = 0$$

we observe that

$$\mathbf{x}'' \sim \mathsf{l}_\diamond^{\top}(\mathbf{x}') \boldsymbol{\lambda}''', \tag{7.25}$$

with $\mathbf{x}''$ lying on $\boldsymbol{\lambda}''$ and

$$\mathbf{x}''' \sim \mathsf{l}_\diamond(\mathbf{x}') \boldsymbol{\lambda}'', \tag{7.26}$$

with $\mathbf{x}'''$ lying on $\boldsymbol{\lambda}'''$. These are transfer relations that allow the prediction of a point in image Ψ_2 resp. Ψ_3 given a point in image Ψ_1 and a corresponding line in image Ψ_3 resp. Ψ_2.

Again using the axiator $\mathsf{S}()$ we obtain from these transfer relations the following trilinearities

$$\mathsf{S}(\mathbf{x}'')\mathsf{I}_{\diamond}^{\top}(\mathbf{x}')\boldsymbol{\lambda}''' = \mathbf{0}, \tag{7.27}$$
$$\mathsf{S}(\mathbf{x}''')\mathsf{I}_{\diamond}(\mathbf{x}')\boldsymbol{\lambda}'' = \mathbf{0}. \tag{7.28}$$

The last and most important type of trilinear constraints is again obtained from relation (7.24). Let us consider a point $\mathbf{x}''$ in image Ψ_2 which corresponds to a point $\mathbf{x}'$ in image Ψ_1. Through $\mathbf{x}''$ we can construct three unique lines $\boldsymbol{\sigma}_i''$ by joining $\mathbf{x}''$ with the three canonical points of $\mathbb{P}^2$ (see table 3.1 on page 15)

$$\begin{aligned}
\boldsymbol{\sigma}_1'' &= \begin{pmatrix} 0, & -1, & y_2 \end{pmatrix}^{\top} \\
\boldsymbol{\sigma}_2'' &= \begin{pmatrix} 1, & 0, & -x_2 \end{pmatrix}^{\top} \\
\boldsymbol{\sigma}_3'' &= \begin{pmatrix} -y_2, & x_2, & 0 \end{pmatrix}^{\top}.
\end{aligned}$$

The line $\boldsymbol{\sigma}_1''$ is parallel to the x''-axis, $\boldsymbol{\sigma}_2''$ is parallel to the y''-axis and $\boldsymbol{\sigma}_3''$ is the connection with the origin; see figure 7.7. These three lines can be used to form the rows of a matrix which turns out to be the axiator $\mathsf{S}(\mathbf{x}'')$. In the same way we can construct three lines $\boldsymbol{\sigma}_j'''$ through the point $\mathbf{x}'''$ in image Ψ_3, which are the rows of the axiator $\mathsf{S}(\mathbf{x}''')$.

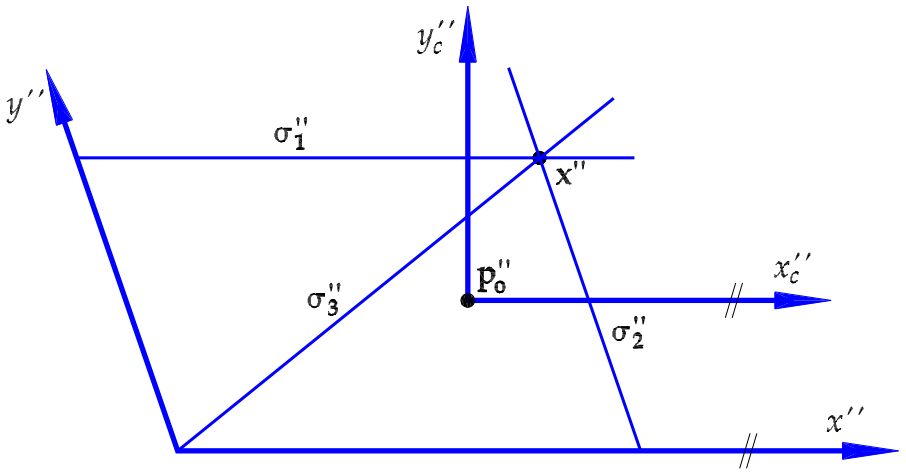

Figure 7.7: *The lines $\boldsymbol{\sigma}_i''$ in image Ψ_2 connecting $\mathbf{x}''$ with the canonical points of $\mathbb{P}^2$.*

Any line $\boldsymbol{\sigma}_i''$ can be used as the line $\boldsymbol{\lambda}''$ in (7.24) and any line $\boldsymbol{\sigma}_j'''$ can be used as the line $\boldsymbol{\lambda}'''$. This gives nine constraints in total, which can be represented as

$$\mathsf{S}(\mathbf{x}''')\mathsf{I}_{\diamond}(\mathbf{x}')\mathsf{S}^{\top}(\mathbf{x}'') = \mathsf{O}. \tag{7.29}$$

These constraints involve a triple of corresponding points and are called **point-trilinearites**; e.g. [Shashua and Wolf 1994]. Because the axiator matrices have rank-2, only 4 of the equations in 7.29 are linearly independent.

So in total we found five kinds of trilinearities, which are summarized in table 7.5 on the following page. Note, however, that each set of the trilinearities (7.23), (7.27), (7.28) and (7.29) is linearly dependent due to the involvement of the axiator $\mathsf{S}()$. This would led to an unnecessary additional computational effort when working with these trilinearities. Therefore it advisable to replace $\mathsf{S}()$ with the reduced version $\mathsf{S}^{red}()$ (see section 3.6). As long as the corresponding

image points do not lie at infinity we can safely assume the third projective coordinate to be non-zero and use $\mathsf{S}^{red}()$ from relation 3.48 on page 28.

For image lines no such general reduction can be applied, as any component can be zero, and therefore each $\mathsf{S}^{red}()$ for lines has to be constructed individually. Another possibility proposed in [Hartley 1994a] is to represent the line in image Ψ_1 by its endpoints $\mathbf{x}'$ and $\mathbf{y}'$ and to use the point-line-line trilinearity (7.24) two times.

trilinearities		
entities	equations	n
$\boldsymbol{\lambda}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''$	$\mathsf{S}^{red}(\boldsymbol{\lambda}') \begin{pmatrix} \boldsymbol{\lambda}'''^{\top} \mathsf{I}_1 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{I}_2 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{I}_3 \boldsymbol{\lambda}'' \end{pmatrix} = \mathbf{0}$	2
$\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''$	$\boldsymbol{\lambda}'''^{\top} \mathsf{I}_{\diamond}(\mathbf{x}') \boldsymbol{\lambda}'' = 0$	1
$\mathbf{x}', \mathbf{x}'', \boldsymbol{\lambda}'''$	$\mathsf{S}^{red}(\mathbf{x}'') \mathsf{I}_{\diamond}^{\top}(\mathbf{x}') \boldsymbol{\lambda}''' = \mathbf{0}$	2
$\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}'''$	$\mathsf{S}^{red}(\mathbf{x}''') \mathsf{I}_{\diamond}(\mathbf{x}') \boldsymbol{\lambda}'' = \mathbf{0}$	2
$\mathbf{x}', \mathbf{x}'', \mathbf{x}'''$	$\mathsf{S}^{red}(\mathbf{x}''') \mathsf{I}_{\diamond}(\mathbf{x}') \mathsf{S}^{red\top}(\mathbf{x}'') = \mathsf{0}$	4

Table 7.5: *The trilinearities of the trifocal tensor. The reduced axiator $\mathsf{S}^{red}(\mathbf{x})$ for points is given by relation 3.48. The last column for the trilinearities gives the number n of linearly independent equations within each group of trilinearities. $\mathsf{I}_{\diamond}(\mathbf{x}')$ is defined in proposition 7.2.*

The line- and point-trilinearites are used when the trifocal tensor is determined for a given set of point and line correspondences. Since the trilinearites are linear in the 27 elements of the tensor they allow – in principle – a direct solution for the trifocal tensor; see section 8.1. For each triple of points and for each triple of lines we get four and two trilinearities respectively. Therefore we need at least 7 triples of points or 13 triples of lines (or a proper combination) to get a linear solution for the tensor.

Note, however, that the tensor has actually 18 degrees of freedom. Therefore 7 triples of points or 13 triples of lines is not the minimum requirement for a solution. Each line trilinearity produces two independent equations, so 9 lines is the minimum requirement. Although this case seems to be disregarded in the literature so far.

For the point trilinearity we encounter a strange schizophrenic behavior: If the 27 elements of the tensor are considered linearly independent, then each point trilinearity produces 4 independent equations, however, if the tensor is considered to have 18 degrees of freedom then each point trilinearity produces only 3 independent equations. Therefore the minimum requirement are 6 points. We will come back to this issue in chapter 8 about the computation of the trifocal tensor.

The missing transfer relations So far we found three kinds of transfer relations: $\{\boldsymbol{\lambda}'', \boldsymbol{\lambda}'''\} \rightarrow \boldsymbol{\lambda}'$ using (7.22), $\{\mathbf{x}', \boldsymbol{\lambda}'''\} \rightarrow \mathbf{x}''$ using (7.25) and $\{\mathbf{x}', \boldsymbol{\lambda}''\} \rightarrow \mathbf{x}'''$ using (7.26). Of course, between three images there exist more transfer configurations. The three presented ones, however, are unique, as they are linear in the elements of the tensor.

Note, that each of these three relations transfers into another image. The missing transfer relations can now be constructed from the given ones; they are, however, no longer linear in the elements of the tensor. In the case of $\{\mathbf{x}'', \mathbf{x}'''\} \to \mathbf{x}'$, $\{\mathbf{x}', \mathbf{x}'''\} \to \mathbf{x}''$ and $\{\mathbf{x}', \mathbf{x}'''\} \to \mathbf{x}''$ the two starting points have to satisfy the epipolar constraint otherwise the transferred point and the two starting points will not give a valid correspondence. All transfer relations are summarized in table 7.6 (with the assumption that the epipolar constraint is satisfied).

transfer relations	
into image Ψ_1	
$\{\boldsymbol{\lambda}'', \boldsymbol{\lambda}'''\} \to \boldsymbol{\lambda}'$	$\boldsymbol{\lambda}' \sim \begin{pmatrix} \boldsymbol{\lambda}'''^{\top} \mathsf{l}_1 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_2 \boldsymbol{\lambda}'' \\ \boldsymbol{\lambda}'''^{\top} \mathsf{l}_3 \boldsymbol{\lambda}'' \end{pmatrix} =: \ell(\boldsymbol{\lambda}'', \boldsymbol{\lambda}''')$
$\{\mathbf{x}'', \boldsymbol{\lambda}'''\} \to \mathbf{x}'$	$\mathbf{x}' \sim \ell(\boldsymbol{\lambda}_1'', \boldsymbol{\lambda}''') \times \ell(\boldsymbol{\lambda}_2'', \boldsymbol{\lambda}''')$ with $\mathbf{x}'' \sim \boldsymbol{\lambda}_1'' \times \boldsymbol{\lambda}_2''$
$\{\boldsymbol{\lambda}'', \mathbf{x}'''\} \to \mathbf{x}'$	$\mathbf{x}' \sim \ell(\boldsymbol{\lambda}'', \boldsymbol{\lambda}_1''') \times \ell(\boldsymbol{\lambda}_2'', \boldsymbol{\lambda}_2''')$ with $\mathbf{x}''' \sim \boldsymbol{\lambda}_1''' \times \boldsymbol{\lambda}_2'''$
$\{\mathbf{x}'', \mathbf{x}'''\} \to \mathbf{x}'$	$\mathbf{x}' \sim \ell(\boldsymbol{\lambda}_1'', \boldsymbol{\lambda}_1''') \times \ell(\boldsymbol{\lambda}_2'', \boldsymbol{\lambda}_1''')$ with $\mathbf{x}'' \sim \boldsymbol{\lambda}_1'' \times \boldsymbol{\lambda}_2''$ and $\mathbf{x}''' \in \boldsymbol{\lambda}_1'''$
into image Ψ_2	
$\{\mathbf{x}', \boldsymbol{\lambda}'''\} \to \mathbf{x}''$	$\mathbf{x}'' \sim \mathsf{l}_\diamond^{\top}(\mathbf{x}') \boldsymbol{\lambda}''' =: \wp_2(\mathbf{x}', \boldsymbol{\lambda}''')$
$\{\mathbf{x}', \mathbf{x}'''\} \to \mathbf{x}''$	$\mathbf{x}'' \sim \wp_2(\mathbf{x}', \boldsymbol{\lambda}_1''')$ with $\mathbf{x}''' \in \boldsymbol{\lambda}_1'''$
$\{\boldsymbol{\lambda}', \boldsymbol{\lambda}'''\} \to \boldsymbol{\lambda}''$	$\boldsymbol{\lambda}'' \sim \wp_2(\mathbf{x}_1', \boldsymbol{\lambda}''') \times \wp_2(\mathbf{x}_2', \boldsymbol{\lambda}''')$ with $\boldsymbol{\lambda}' \sim \mathbf{x}_1' \times \mathbf{x}_2'$
$\{\boldsymbol{\lambda}', \mathbf{x}'''\} \to \mathbf{x}''$	$\mathbf{x}'' \sim (\wp_2(\mathbf{x}_1', \boldsymbol{\lambda}_1''') \times \wp_2(\mathbf{x}_2', \boldsymbol{\lambda}_1''')) \times (\wp_2(\mathbf{x}_1', \boldsymbol{\lambda}_2''') \times \wp_2(\mathbf{x}_2', \boldsymbol{\lambda}_2'''))$ with $\boldsymbol{\lambda}' \sim \mathbf{x}_1' \times \mathbf{x}_2'$ and $\mathbf{x}''' \sim \boldsymbol{\lambda}_1''' \times \boldsymbol{\lambda}_2'''$
into image Ψ_3	
$\{\mathbf{x}', \boldsymbol{\lambda}''\} \to \mathbf{x}'''$	$\mathbf{x}''' \sim \mathsf{l}_\diamond(\mathbf{x}') \boldsymbol{\lambda}'' =: \wp_3(\mathbf{x}', \boldsymbol{\lambda}'')$
$\{\mathbf{x}', \mathbf{x}''\} \to \mathbf{x}'''$	$\mathbf{x}''' \sim \wp_3(\mathbf{x}', \boldsymbol{\lambda}_1'')$ with $\mathbf{x}'' \in \boldsymbol{\lambda}_1''$
$\{\boldsymbol{\lambda}', \boldsymbol{\lambda}''\} \to \boldsymbol{\lambda}'''$	$\boldsymbol{\lambda}''' \sim \wp_3(\mathbf{x}_1', \boldsymbol{\lambda}'') \times \wp_3(\mathbf{x}_2', \boldsymbol{\lambda}'')$ with $\boldsymbol{\lambda}' \sim \mathbf{x}_1' \times \mathbf{x}_2'$
$\{\boldsymbol{\lambda}', \mathbf{x}''\} \to \mathbf{x}'''$	$\mathbf{x}''' \sim (\wp_3(\mathbf{x}_1', \boldsymbol{\lambda}_1'') \times \wp_3(\mathbf{x}_2', \boldsymbol{\lambda}_1'')) \times (\wp_3(\mathbf{x}_1', \boldsymbol{\lambda}_2'') \times \wp_3(\mathbf{x}_2', \boldsymbol{\lambda}_2''))$ with $\boldsymbol{\lambda}' \sim \mathbf{x}_1' \times \mathbf{x}_2'$ and $\mathbf{x}'' \sim \boldsymbol{\lambda}_1'' \times \boldsymbol{\lambda}_2''$

Table 7.6: *All possible transfer relations involving the trifocal tensor.* $\mathsf{l}_\diamond(\mathbf{x}')$ *is defined in proposition 7.2 on page 67.*

The transfer relations are of interest if the tensor is already known, computed from image correspondences or from given projection matrices, and can be used to transfer the content of two source images into the other one (the target image). If the target image is a *real* image, the trans-

ferred position of features can be used to initialize a matching technique in the target image. If the target image is *virtual* with a given orientation, then the transfer relations can be used to form the content of this new image by transferring pixel by pixel. This process is called *novel view synthesis* or *image formation*. Reports on this transfer topic can be found e.g. in [Avidan and Shashua 1997] or [Mayer 2002].

Degenerate transfer configurations. The line transfer (7.22) into the first image fails in the following cases; adapted from [Faugeras and Luong 2001] and [Hartley and Zisserman 2001]:

(a) $\boldsymbol{\lambda}''$and $\boldsymbol{\lambda}'''$are epipolar lines with respect to the first image. In this case the corresponding 3D line is a projection ray of the first image.

(b) $\boldsymbol{\lambda}''$and $\boldsymbol{\lambda}'''$are epipolar lines with respect to the image pair Ψ_2 and Ψ_3. In this case the projection plane of $\boldsymbol{\lambda}''$and $\boldsymbol{\lambda}'''$are coplanar and therefore do not intersect in space. Since in case of $\mathbf{Z}_2 = \mathbf{Z}_3$ the coplanarity of the projection planes happens for all 3D lines, the following two special cases exist.

   (b1) The projection centers $\mathbf{Z}_2$ and $\mathbf{Z}_3$ coincide. In this case the projection planes of $\boldsymbol{\lambda}''$and $\boldsymbol{\lambda}'''$coincide, and the images Ψ_2 and Ψ_3 are related by H_∞. The relation between image Ψ_1 and any of the other two images is then described by the fundamental matrix. Consequently the line in image Ψ_1 is *not uniquely defined* by the corresponding pair of lines in image Ψ_2 *and* image Ψ_3.

   (b2) The three projection centers coincide ($\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{Z}_3$). In this case the trifocal tensor is the zero tensor and any pair of images is related by the homography H_∞ induced by the plane at infinity. Consequently the line in image Ψ_1 is already defined by the corresponding line in image Ψ_2 *or* image Ψ_3.

A special degeneracy of the both cases (a) and (b) are all 3D lines lying in the **trifocal plane**, which is the plane passing through the projection centers of all three images.

The point transfer relation (7.25) into the second image fails in the following cases

(a) The line $\boldsymbol{\lambda}'''$is the epipolar line of $\mathbf{x}'$. In this case the projection plane of $\boldsymbol{\lambda}'''$and the projection ray of $\mathbf{x}'$are incident, however, this case can be avoided by choosing another corresponding line $\boldsymbol{\lambda}'''$through $\mathbf{x}'''$.

(b) The corresponding object point lies on the basis of the first and third image. Since in case of $\mathbf{Z}_1 = \mathbf{Z}_3$ all object points lie on a line through $\mathbf{Z}_1$ and $\mathbf{Z}_3$, the following two special cases exist.

   (b1) The projection centers $\mathbf{Z}_1$ and $\mathbf{Z}_3$ coincide. In this case the projection ray of $\mathbf{x}'$ lies in the projection plane of $\boldsymbol{\lambda}''$and the images Ψ_1 and Ψ_3 are related by H_∞. The relation between image Ψ_2 and any of the other two images is then described by the fundamental matrix. Consequently the point in image Ψ_1 is *not uniquely defined* by the point in image Ψ_1 *and* the line in image Ψ_3.

   (b2) The three projection centers coincide ($\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{Z}_3$). In this case the trifocal tensor is the zero tensor and any pair of images is related by the homography H_∞ induced by the plane at infinity. Consequently the point in image Ψ_2 is already defined by the corresponding point in image Ψ_1 *or* image Ψ_3.

Analogous cases exist for the point transfer relation (7.26) into the third image. Note that the point transfer between the three images could also be performed using the three fundamental matrices of the three views. This so-called *epipolar transfer* can not handle lines and fails for image points if the three projection centers and the corresponding object point are coplanar. For more details on these transfer relations see e.g. [Hartley and Zisserman 2001].

7.4 Retrieving the orientation parameters

Assume we are given the trifocal tensor, e.g. by determining it with the trilinearities and a set of corresponding features (see section 8). Then we can extract the epipoles $\mathbf{e}'_2$, $\mathbf{e}'_3$, $\mathbf{e}''_1$, $\mathbf{e}''_3$, $\mathbf{e}'''_1$ and $\mathbf{e}'''_2$, the fundamental matrices F_{12}, F_{13} and F_{23}, and the projection matrices P_1, P_2 and P_3 out of the tensor. The latter can not be retrieved uniquely. Since the trifocal tensor describes only the relative orientation of the three images the projection matrices can be retrieved only up to a certain transformation in space. The type of this transformation depends on the knowledge of the interior orientation of the images and is a general projective transformation in $\mathbb{P}^3$ if it is unknown, and a spatial similarity transformation in $\mathbb{R}^3$ if it is known. If the interior orientation of the images is *unknown but common to all three images*, then this common interior orientation can be retrieved in general; see section 7.4.4.

For this chapter we will assume that the three images have distinct projection centers, because only in this case the existence of all epipoles and fundamental matrices is defined.

7.4.1 Retrieving the epipoles

The six epipoles of the tree images can be extracted easily from the tensor using the tensorial slices I_i, J_j and K_k.

The epipoles $\mathbf{e}''_1$ and $\mathbf{e}'''_1$. From proposition 7.4 on page 69 (about the properties of rank-2 correlation slices I_i) we see that these two epipoles arise generally as the left nullspaces of the matrices N_R and N_L, which itself are made up of the left resp. right nullspaces of the correlation slices. This method is proposed in many articles e.g. [Hartley 1994b], [Faugeras and Luong 2001] or [Hartley and Zisserman 2001]; but it is only applicable if all three correlation slices have maximum rank-2. For I_i matrices with smaller rank the two epipoles can nevertheless be computed, however, the algorithm can not use the two N matrices, because they are not defined. This is due to the fact, that the columns of the two N matrices are supposed to be made up of the single null spaces of the I_i matrices, however, in case of rank-1 or rank-0 correlation slices the nullspace is a pencil of lines or all lines in the image. Therefore, the algorithm for retrieving the two epipoles must consider the ranks of the three I_i matrices and choose the respective methods from table 7.1 on page 68. This situation is, of course, a mayor drawback, as the rank-determination of matrices is an ill-posed problem in case of noisy data.

Another method, which in general returns these two epipoles, is to compute the general eigenvalues of $\mathsf{J}_p - \kappa\mathsf{J}_q$ resp. $\mathsf{K}_p - \kappa\mathsf{K}_q$. The general eigenvalue with multiplicity 2 is the ratio of the pth and qth coordinate of epipole $\mathbf{e}''_1$ resp. $\mathbf{e}'''_1$; see propositions 7.8 and 7.12 on page 73 and 76 respectively. This method was presented in [Shashua and Werman 1995], however, it suffers from the same drawback as the previous method, that the ranks of the homography slices

need to be considered. Especially it is the matrix J_q resp. K_q (the one the general eigenvalue is associated with) which needs to be regular; the rank of the partner matrix is not important. A workaround would be to select the J resp. K matrix with the best condition number for this important matrix of the general eigenvalue problem. However, from table 7.4 on page 77 we see that in principle all three J or K matrices can be singular, thus a certain insecurity remains.

Note that the singularity of the tensor slices depends on the incidence of the projection centers with the principal rays and planes and that the orientation of the principal rays and planes of an image if defined by the 3D image coordinate system (see figure 5.3 on page 38), which itself depends on the image measurement system. Therefore we could avoid the singularity of the tensor slices by adapting the image coordinate systems of the three views. This is, however, a bad strategy as we do not know in advance if our current image systems are appropriate or not and if not how to change them (and not to risk to produce new singular matrices). Therefore it is better to aim for methods for retrieving the epipoles, that work independent on the choice of the image coordinate system and thus independent on the ranks of the tensor slices.

One such method was proposed by Spetsakis and Aloimonos already in 1990, [Spetsakis and Aloimonos 1990], although they actually did not deal with the trifocal tensor as such but considered three matrices that correspond to the correlation slices (for calibrated images). Their method is in principle identical to the first method presented above, however, the authors were aware of the rank problem and solved it in the following way.

In proposition 7.2 on page 67 we introduced the linear combination $\mathsf{l}_\diamond$ of all three correlation slices

$$\mathsf{l}_\diamond(\mathbf{p}') = u'\mathsf{l}_1 + v'\mathsf{l}_2 + w'\mathsf{l}_3,$$

which is the dual line correlation induced by the projection ray of the point $\mathbf{p}' \sim (u', v', w')^\top$ in image Ψ_1. $\mathsf{l}_\diamond(\mathbf{p}')$ is therefore of rank-2 in general and its right nullspace $\boldsymbol{\eta}(\mathbf{p}')$ passes through the epipole $\mathbf{e}_1''$. Let $\mathsf{N}(\mathbf{p}') = \mathrm{adj}(\mathsf{l}_\diamond(\mathbf{p}'))$ denote the adjoint (defined in (B.14) on page 179) of $\mathsf{l}_\diamond(\mathbf{p}')$. Then $\mathsf{N}(\mathbf{p}')$ is of rank 1 in general (and of rank 0 if $\mathrm{rank}(\mathsf{l}_\diamond(\mathbf{p}')) < 2$) and all of its column vectors are proportional to $\boldsymbol{\eta}(\mathbf{p}')$. Therefore we could choose any column $\mathsf{N}\boldsymbol{i}_c$, $c \in \{1, 2, 3\}$, and form one defining equation

$$\boldsymbol{i}_c{}^\top \mathsf{N}(\mathbf{p}')^\top \mathbf{e}_1'' = 0.$$

However, since we do not know in advance which column c is best to choose (because in principle the chosen column can be the zero vector), it is better to choose all three and combine them in a least-squares way

$$\mathbf{e}_1''{}^\top \mathsf{N}(\mathbf{p}')\mathsf{N}(\mathbf{p}')^\top \mathbf{e}_1'' \to \min. \tag{7.30}$$

If the correlation slices have rank ≤ 2 the minimum will be zero[3]. This relation can be used for *any* point $\mathbf{p}'$, even if the respective matrix $\mathsf{l}_\diamond(\mathbf{p}')$ has not rank-2, since in this case $\mathsf{N}(\mathbf{p}')$ is the zero matrix. For the determination of $\mathbf{e}_1''$ we need a second equation, which is obtained in the same way but for another point $\mathbf{p}'$. Particularly if we choose $\mathbf{p}' = \boldsymbol{i}_i$, then $\mathsf{l}_\diamond(\mathbf{p}')$ turns into the correlation slice l_i. However, since we can not be sure if the matrices $\mathsf{l}_\diamond(\mathbf{p}')$ obtained for

[3] If the trifocal tensor is computed from noisy data without considering the internal constraints, then the correlation slices will be of rank-3 in general and the minimum will be ≥ 0.

two chosen points $\mathbf{p}'$are not degenerate, it is better to choose all points $\mathbf{p}'$and to integrate (7.30) over the unit sphere with respect to $\mathbf{p}'$. Doing this we then have to solve

$$\mathbf{e}_1''^{\top} \mathsf{M} \mathbf{e}_1'' \rightarrow \min,$$

with

$$\mathsf{M} = \int_{\text{unit sphere}} \mathsf{N}(\mathbf{p}')\mathsf{N}(\mathbf{p}')^{\top} d\mathbf{p}'.$$

This integral can be solved explicitly (see [Spetsakis and Aloimonos 1990]) and we get

$$\begin{aligned} \mathsf{M} = 3\Big(\mathrm{adj}(\mathsf{Z}_1)\mathrm{adj}(\mathsf{Z}_1)^{\top} + \mathrm{adj}(\mathsf{Z}_2)\mathrm{adj}(\mathsf{Z}_2)^{\top} + \mathrm{adj}(\mathsf{Z}_3)\mathrm{adj}(\mathsf{Z}_3)^{\top}\Big) + \\ \mathrm{adj}(\mathsf{Z}_1)\mathrm{adj}(\mathsf{Z}_2)^{\top} + \mathrm{adj}(\mathsf{Z}_1)\mathrm{adj}(\mathsf{Z}_3)^{\top} + \mathrm{adj}(\mathsf{Z}_2)\mathrm{adj}(\mathsf{Z}_1)^{\top} + \mathrm{adj}(\mathsf{Z}_2)\mathrm{adj}(\mathsf{Z}_3)^{\top} + \\ \mathrm{adj}(\mathsf{Z}_3)\mathrm{adj}(\mathsf{Z}_1)^{\top} + \mathrm{adj}(\mathsf{Z}_3)\mathrm{adj}(\mathsf{Z}_2)^{\top} + \mathsf{A}_{12}\mathsf{A}_{12}{}^{\top} + \mathsf{A}_{13}\mathsf{A}_{13}{}^{\top} + \mathsf{A}_{23}\mathsf{A}_{23}{}^{\top}, \end{aligned} \tag{7.31}$$

with

$$\begin{aligned} \mathsf{A}_{12} &= \mathrm{adj}(\mathsf{Z}_1 + \mathsf{Z}_2) - \mathrm{adj}(\mathsf{Z}_1) - \mathrm{adj}(\mathsf{Z}_2), \\ \mathsf{A}_{13} &= \mathrm{adj}(\mathsf{Z}_1 + \mathsf{Z}_3) - \mathrm{adj}(\mathsf{Z}_1) - \mathrm{adj}(\mathsf{Z}_3), \\ \mathsf{A}_{23} &= \mathrm{adj}(\mathsf{Z}_2 + \mathsf{Z}_3) - \mathrm{adj}(\mathsf{Z}_2) - \mathrm{adj}(\mathsf{Z}_3)\ . \end{aligned}$$

Since we will need this matrix M also for retrieving other epipoles, we used general argument-matrices Z_i. For the retrieving the epipole $\mathbf{e}_1''$ we have to set $\mathsf{Z}_i = \mathsf{I}_i$.

The matrix M will have rank-2 in general and $\mathbf{e}_1''$ is obtained as the eigenvector of the *smallest* eigenvalue of M. This solution is valid, as long as none of the other two projection centers coincides with the projection center of the first image. Otherwise all three correlation slices I_i have rank ≤ 1 (with $\mathsf{I}_1 \sim \mathsf{I}_2 \sim \mathsf{I}_3$, see proposition 7.3 on page 68) and therefore M would be of rank-0.

The epipole $\mathbf{e}_1'''$ is found in exactly the same way but by using $\mathsf{Z}_i = \mathsf{I}_i{}^{\top}$ in equation (7.31).

The epipoles $\mathbf{e}_2'$ and $\mathbf{e}_3'$. After having obtained the epipoles $\mathbf{e}_1''$ and $\mathbf{e}_1'''$, the remaining ones can be found using the homography matrices. From relation (7.16) we see that $\mathbf{e}_3'$ can be computed as

$$\mathbf{e}_3' \sim \mathsf{J}_j{}^{-1}\mathbf{e}_1'''.$$

For this computation only one of the three homography matrices J_j is required and to avoid the inversion of a nearly singular matrix the one with the best condition number should be chosen. However, from table 7.4 on page 77 we see that it may happen that all three J_j are singular. In this case or in the vicinity to it, this method is unreliable.

Actually the inversion of J_j is not required, because by applying the identity check in $\mathbb{P}^2$ (section 3.6) onto the relation (7.16) we simply get

$$\mathsf{S}(\mathbf{e}_1''')\mathsf{J}_j\mathbf{e}_3' = \mathbf{0}, \tag{7.32}$$

and can compute $\mathbf{e}_3'$ as the nullspace of the matrix $\mathsf{S}(\mathbf{e}_1''')\mathsf{J}_j$, which is actually the fundamental matrix $\mathsf{F}_{13}{}^{\top}$. This method is valid as long as the chosen matrix J_j has at least rank-2 and from

table 7.4 we see that for all possible cases at least one such matrix is available. However, to avoid choosing a particular J_j matrix, we can again use the method proposed by Spetsakis and Aloimonos and build up the matrix M in (7.31) using the matrices $\mathsf{Z}_j = \mathsf{S}(\mathbf{e}_1''')\mathsf{J}_j$. In this case the epipole $\mathbf{e}_3'$ is obtained as the eigenvector to the *largest* eigenvalue of the matrix M.

The epipole $\mathbf{e}_2'$ is found in exactly the same way but by considering the matrices $\mathsf{Z}_k = \mathsf{S}(\mathbf{e}_1'')\mathsf{K}_k$.

The epipoles $\mathbf{e}_3''$ and $\mathbf{e}_2'''$. The epipole $\mathbf{e}_2'''$ could be computed for example by

$$\mathbf{e}_2''' \sim \mathsf{J}_j \mathbf{e}_2'.$$

This relation follows from the fact that the homography planes associated with the three J matrices are the principal planes of image Ψ_2, which intersect in the projection center $\mathbf{Z}_2$. Only one J matrix is required to find $\mathbf{e}_2'''$ and its rank need to be at least 2. Thus we again have the problem to choose a proper one.

Another method, that avoids this problem, is presented in [Lasenby and Lasenby 1998]. The epipole $\mathbf{e}_2'''$ is the eigenvector of the largest eigenvalue of the matrix $\mathsf{l}_\diamond(\mathbf{e}_2')$

$$\mathsf{l}_\diamond(\mathbf{e}_2') = u'\mathsf{l}_1 + v'\mathsf{l}_2 + w'\mathsf{l}_3, \tag{7.33}$$

with $\mathbf{e}_2' \sim (u', v', w')^\top$.

The matrix $\mathsf{l}_\diamond(\mathbf{e}_2')$ represents the dual line correlation from image Ψ_2 to image Ψ_3 induced by the projection ray of the epipole $\mathbf{e}_2'$ in image Ψ_1. Since this ray is actually the basis between image Ψ_1 and image Ψ_2 it contains the projection center of the source image Ψ_2 and therefore $\mathsf{l}_\diamond(\mathbf{e}_2')$ has rank-1 (provided $\mathbf{Z}_1 \neq \mathbf{Z}_2$) and all its columns are proportional to $\mathbf{e}_2'''$; cf. table 7.1 on page 68. And therefore we actually need not compute the eigenvectors of $\mathsf{l}_\diamond(\mathbf{e}_2')$ and can simply choose the column with largest norm to obtain $\mathbf{e}_2'''$.

The epipole $\mathbf{e}_3''$ can be found in the same way but by considering the matrix $\mathsf{l}_\diamond^\top(\mathbf{e}_3')$.

Summary of retrieving the epipoles

Table 7.7 on the next page gives a summary on retrieving the six epipoles in the three images using the trifocal tensor.

7.4.2 Retrieving the fundamental matrices

The fundamental matrices can be derived from the trifocal tensor basically in two ways. The first method requires the epipoles to be computed prior and the second method does not.

epipole	method
$\mathbf{e}_1''$	the eigenvector of the *smallest* eigenvalue of matrix M set up using the matrices $\mathsf{Z}_i = \mathsf{I}_i$
$\mathbf{e}_1'''$	the eigenvector of the *smallest* eigenvalue of matrix M set up using the matrices $\mathsf{Z}_i = \mathsf{I}_i^\top$
$\mathbf{e}_2'$	the eigenvector of the *largest* eigenvalue of matrix M set up using the matrices $\mathsf{Z}_k = \mathsf{S}(\mathbf{e}_1'')\mathsf{K}_k$
$\mathbf{e}_3'$	the eigenvector of the *largest* eigenvalue of matrix M set up using the matrices $\mathsf{Z}_j = \mathsf{S}(\mathbf{e}_1''')\mathsf{J}_j$
$\mathbf{e}_3''$	the column of matrix $\mathsf{I}_\diamond^\top(\mathbf{e}_3')$ with *largest* norm
$\mathbf{e}_2'''$	the column of matrix $\mathsf{I}_\diamond(\mathbf{e}_2')$ with *largest* norm

Table 7.7: *Summary of retrieving the epipoles using the trifocal tensor. The matrix M is defined by relation* (7.31) *on page 85 and the matrix $\mathsf{I}_\diamond(\cdot)$ is defined by relation* (7.12) *on page 67.*

Using the epipoles

With the epipoles computed in the previous section and the homography slices the fundamental matrices between the three pairs of images can be computed using proposition 6.9 on page 56, which describes the relation between plane homographies and fundamental matrices. The fundamental matrix F_{13} between image Ψ_1 and Ψ_3 is found by

$$\mathsf{F}_{13} \sim \mathsf{J}_x{}^\top \mathsf{S}(\mathbf{e}_1''') . \tag{7.34}$$

For this computation only one (denoted J_x) of the three J_j matrices is needed, which must have at least rank-2. From table 7.4 on page 77 we see that one such matrix is always available – but how to choose the best one? From proposition 7.7 on page 72 we see that $\operatorname{rank}(\mathsf{J}_j) = 1$ is equivalent to $e_1''^{(j)} = 0$ and therefore the J matrix for the index of the maximum absolute element of the epipole $\mathbf{e}_1''$ is the one that is farthest away from having rank ≤ 1. Another choice for J_x, which avoids this maximum search, is to use the matrix $\mathsf{J}_\diamond(\mathbf{e}_1'')$ – the linear combination of the J homography slices introduced in proposition 7.6 on page 71:

$$\mathsf{J}_\diamond(\mathbf{e}_1'') = u''\mathsf{J}_1 + v''\mathsf{J}_2 + w''\mathsf{J}_3, \tag{7.35}$$

with $\mathbf{e}_1'' \sim (u'', v'', w'')^\top$. The matrix $\mathsf{J}_\diamond(\mathbf{e}_1'')$ represents the homography induced by the projection plane of a line $\boldsymbol{\lambda_2} = \mathbf{e}_1''^*$ in Ψ_2 whose coordinates are given by $\boldsymbol{\lambda_2} \sim \mathbf{e}_1''$; i.e. the dual of $\mathbf{e}_1''$. This choice ensures that the homography plane does not contain the projection center $\mathbf{Z}_1$ and thus has certainly rank > 1.

Note that F_{13} can also be expressed in terms of the correlation slices as

$$\mathsf{F}_{13} \sim \begin{bmatrix} \mathbf{e}_1''^\top \mathsf{I}_1{}^\top \\ \mathbf{e}_1''^\top \mathsf{I}_2{}^\top \\ \mathbf{e}_1''^\top \mathsf{I}_3{}^\top \end{bmatrix} \mathsf{S}(\mathbf{e}_1''') ,$$

as this expression is identical to using $\mathsf{J}_\diamond(\mathbf{e}_1'')$ in equation (7.34). The geometric explanation is that we build up F_{13} row-wise (cf. proposition 6.3 on page 47) using the left nullspaces of the correlation slices. These left nullspaces are the maps of the principal rays $\boldsymbol{\mathcal{R}}_i'$ of image Ψ_1 into image Ψ_3, which can be expressed as $\mathsf{S}(\mathbf{e}_1''')\mathsf{l}_i\boldsymbol{\lambda}''$. The line $\boldsymbol{\lambda}''$ can be any line not going through the epipole $\mathbf{e}_1''$, otherwise $\mathsf{l}_i\boldsymbol{\lambda}'' \sim \mathbf{e}_1'''$. To ensure $\mathbf{e}_1'' \notin \boldsymbol{\lambda}''$ we can use the dual of the point $\mathbf{e}_1''$ as the line $\boldsymbol{\lambda}''$.

Analogously the fundamental matrix F_{12} between image Ψ_1 and Ψ_2 is found by

$$\mathsf{F}_{12} \sim \mathsf{K}_x{}^\top \mathsf{S}(\mathbf{e}_1''), \tag{7.36}$$

where K_x is chosen for the index of the maximum absolute element of the epipole $\mathbf{e}_1'''$ or as

$$\mathsf{K}_\diamond(\mathbf{e}_1''') = u'''\mathsf{K}_1 + v'''\mathsf{K}_2 + w'''\mathsf{K}_3,$$

with $\mathbf{e}_1''' \sim (u''', v''', w''')^\top$. And another representation would be

$$\mathsf{F}_{12} \sim \begin{bmatrix} \mathbf{e}_1'''^\top \mathsf{l}_1{}^\top \\ \mathbf{e}_1'''^\top \mathsf{l}_2{}^\top \\ \mathbf{e}_1'''^\top \mathsf{l}_3{}^\top \end{bmatrix} \mathsf{S}(\mathbf{e}_1'').$$

For determining the fundamental matrix F_{23} between the images Ψ_2 and Ψ_3 we need a homography between these two images. None of the three different slices of the trifocal tensor $T(\Psi_1, \Psi_2, \Psi_3)$ represents such a homography directly. But it is nevertheless possible to derive a homography between the images Ψ_2 and Ψ_3 using the J and K homography slices – provided all three projection centers are distinct. Although this homography has only rank-2, it is sufficient to derive the third fundamental matrix.

Proposition 7.13 (A rank-2 homography between the images Ψ_2 and Ψ_3.) *Using the* J *and* K *homography slices of the trifocal tensor $T(\Psi_1, \Psi_2, \Psi_3)$, with distinct projection centers, a rank-2 homography from image Ψ_2 to image Ψ_3 can be computed as*

$$\mathsf{H}_{23} \sim \mathsf{J}_\diamond(\boldsymbol{\lambda}'')\mathsf{K}_\diamond{}^{-1}(\boldsymbol{\mu}''') - \frac{1}{\boldsymbol{\mu}'''^\top \mathbf{e}_1'''} \mathbf{e}_1''' \boldsymbol{\lambda}''^\top,$$

with $\boldsymbol{\lambda}'' = \mathbf{e}_1''^*$, the dual *of $\mathbf{e}_1''$, and $\boldsymbol{\mu}''' = 1/2(\mathbf{e}_1''' + \mathbf{e}_2''')^*$. This homography is induced by the projection plane of the line $\boldsymbol{\mu}'''$ in image Ψ_3. $\mathsf{J}_\diamond$ and $\mathsf{K}_\diamond$ follow from* (7.15) *and* (7.20) *on page 71 and 75 respectively.*

Analogously, the rank-2 homography from image Ψ_3 to Ψ_2 induced by the line $\boldsymbol{\lambda}''$ in image Ψ_2 can be computed as

$$\mathsf{H}_{32} \sim \mathsf{K}_\diamond(\boldsymbol{\mu}''')\mathsf{J}_\diamond{}^{-1}(\boldsymbol{\lambda}'') - \frac{1}{\boldsymbol{\lambda}''^\top \mathbf{e}_1''} \mathbf{e}_1'' \boldsymbol{\mu}'''^\top,$$

with $\boldsymbol{\lambda}'' = 1/2(\mathbf{e}_1'' + \mathbf{e}_3'')^*$ and $\boldsymbol{\mu}''' = \mathbf{e}_1'''^*$. $\mathsf{J}_\diamond$ and $\mathsf{K}_\diamond$ follow again from (7.15) *and* (7.20).

Proof:
We will only give the proof for the expression of H_{23}, as for the proof of H_{32} only the roles

of the images Ψ_2 and Ψ_3 need to be interchanged. Using the representations for the linear combinations $\mathsf{J}_\diamond$ and $\mathsf{K}_\diamond$ from proposition 7.6 and 7.10 we get

$$\mathsf{J}_\diamond(\boldsymbol{\lambda}'') = \mathsf{M}_3\Big(\boldsymbol{\lambda}''^\top \mathbf{e}_1'' \mathsf{I} + (\mathsf{Z}_3 - \mathsf{Z}_1)\boldsymbol{\lambda}''^\top \mathsf{M}_2\Big)\mathsf{M}_1{}^{-1},$$
$$\mathsf{K}_\diamond{}^{-1}(\boldsymbol{\mu}''') = -\frac{1}{\boldsymbol{\mu}'''^\top \mathbf{e}_1'''}\mathsf{M}_1\Big(\mathsf{I} - \frac{1}{\boldsymbol{\mu}'''^\top \mathbf{e}_2'''}(\mathsf{Z}_2 - \mathsf{Z}_1)\boldsymbol{\mu}'''^\top \mathsf{M}_3\Big)\mathsf{M}_2{}^{-1},$$

where proposition B.1 together with the epipole relations (7.2) and (7.3) was used for the inverse of $\mathsf{K}_\diamond{}^{-1}(\boldsymbol{\mu}''')$. The multiplication of these two matrices gives

$$\mathsf{J}_\diamond(\boldsymbol{\lambda}'')\mathsf{K}_\diamond{}^{-1}(\boldsymbol{\mu}''') = \underbrace{\frac{-\boldsymbol{\lambda}''^\top \mathbf{e}_1''}{\boldsymbol{\mu}'''^\top \mathbf{e}_1''' \cdot \boldsymbol{\mu}'''^\top \mathbf{e}_2'''}}_{t}\underbrace{\mathsf{M}_3\Big(\boldsymbol{\mu}'''^\top \mathbf{e}_2''' \mathsf{I} + (\mathsf{Z}_3 - \mathsf{Z}_2)\boldsymbol{\mu}'''^\top \mathsf{M}_3\Big)\mathsf{M}_2{}^{-1}}_{\mathsf{H}_{23}} + \frac{1}{\boldsymbol{\mu}'''^\top \mathbf{e}_1'''}\mathbf{e}_1'''\boldsymbol{\lambda}''^\top.$$

The comparison of the underbraced term H_{23} with the homography proposition 6.7 shows that it is a homography from image Ψ_2 to image Ψ_3 induced by a plane with normal vector $\mathsf{M}_3{}^\top\boldsymbol{\mu}'''$ going through $\mathbf{Z}_3$; i.e. the projection plane $\boldsymbol{\Sigma}(\boldsymbol{\mu}''')$ of image Ψ_3. Therefore the rank of H_{23} is 2 in general. All three scalars in t must be $\neq 0$ therefore: $\boldsymbol{\lambda}''^\top \mathbf{e}_1'' \neq 0 \rightarrow \boldsymbol{\lambda}'' = \mathbf{e}_1''^*$, this ensures that $\mathsf{J}_\diamond(\boldsymbol{\lambda}'')$ has at least rank-2; $\boldsymbol{\mu}'''^\top \mathbf{e}_1''' \neq 0$ and $\boldsymbol{\mu}'''^\top \mathbf{e}_2''' \neq 0 \rightarrow \boldsymbol{\mu}''' = 1/2(\mathbf{e}_1''' + \mathbf{e}_2''')^*$, this ensures that $\mathsf{K}_\diamond(\boldsymbol{\mu}''')$ has rank-3. To avoid annulling of $\boldsymbol{\mu}'''$ in case of collinear projection centers, it should be checked that $\mathbf{e}_1'''^\top \mathbf{e}_2''' \geq 0$.

□

Using this proposition the fundamental matrices F_{23} and F_{32} can be found by

$$\begin{aligned}\mathsf{F}_{23} &\sim \mathsf{H}_{23}{}^\top \mathsf{S}(\mathbf{e}_2'''),\\ \mathsf{F}_{32} &\sim \mathsf{H}_{32}{}^\top \mathsf{S}(\mathbf{e}_3'').\end{aligned} \tag{7.37}$$

Without the epipoles

In [Shashua and Werman 1995] a simple method is presented how to computed the fundamental matrices F_{12} and F_{13} from the trifocal tensor without recovering the epipoles at first. From the expressions (6.3) and (7.19) for F_{12} and K_k in the elements of the projection matrices and using the epipole definitions (7.2) and (7.3) together with the axiator relation (B.12) it can be seen easily, that the following two equations hold

$$\begin{aligned}\mathsf{F}_{12}\mathsf{K}_k &= e_1'''^{(k)}/\det(\mathsf{M}_1)\mathsf{S}(\mathbf{e}_2'),\\ \mathsf{K}_k^\top\mathsf{F}_{12}^\top &= -e_1'''^{(k)}/\det(\mathsf{M}_1)\mathsf{S}(\mathbf{e}_2'),\end{aligned}$$

where $e_1'''^{(k)}$ is the k-th coordinate of $\mathbf{e}_1'''$. Therefore the combined relation

$$\mathsf{F}_{12}\mathsf{K}_k + \mathsf{K}_k^\top\mathsf{F}_{12}^\top = \mathsf{0}_3 \tag{7.38}$$

can be used to determine the fundamental matrix F_{12} directly using the K_k homographies. Observe that this symmetric relation produces only 6 independent equations in general. Therefore

the relation (7.38) must be set up for at least two (and best for all three) K_k homographies. This way we get an overdetermined homogenous system of equations $Af = \mathbf{0}$ with

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}$$

where $A_k = \mathsf{K}_k^\top \otimes I_3 + I_3 \otimes \mathsf{K}_k^\top T_3$ (T_3 is defined in relation (B.29)) and $\mathbf{f} = \mathrm{vec}\,(F_{12})$. The solution for $\mathbf{f}$ is obtained as the eigenvector to the smallest eigenvalue of $A^\top A$.

Similar relations allow the determination of F_{13} using the J_j homographies. The epipoles $\mathbf{e}'_2$, $\mathbf{e}'_3$, $\mathbf{e}''_1$ and $\mathbf{e}'''_1$ are then easily obtained from these fundamental matrices using proposition 6.3.

Note that this method can be of disadvantage if the computation time is of concern, because it requires the solution of eigenvalue problems of dimension 9, whereas the methods based on first computing the epipoles always involve systems of dimension 3. Especially if one is only interested in the epipoles then the methods presented in section 7.4.1 are preferred.

7.4.3 Retrieving the projection matrices

In section 7.1 we saw that the trifocal tensor depends only on the relative orientation of the three images. Therefore we could change the object system by any homography in $\mathbb{P}^3$ without effecting the trifocal tensor. If this homography in $\mathbb{P}^3$ is represented by a regular homogenous 4×4 matrix H, then the point projection matrices would accordingly change as $\overline{\mathsf{P}}_h \sim \mathsf{P}_h \mathsf{H}^{-1}$; see equation (5.14). In particular, we can choose H in the way that

$$\mathsf{P}_1 \mathsf{H}^{-1} \sim \overline{\mathsf{P}}_1 = [I_3, \mathbf{0}]\,. \tag{7.39}$$

Note, however, that from this relation H is not uniquely determined, because we only get 12 equations for the 16 elements of H; so 4 degrees of freedom remain. With $\mathsf{P}_1 \sim \mathsf{M}_1\,[I, -\mathbf{Z}_1]$ the general form of H^{-1} is

$$\mathsf{H}^{-1} \sim \begin{bmatrix} {\mathsf{M}_1}^{-1} + \mathbf{Z}_1 \mathbf{v}^\top & w\mathbf{Z}_1 \\ \mathbf{v}^\top & w \end{bmatrix}, \tag{7.40}$$

where $\mathbf{v}^\top$ and w are the remaining 4 degrees of freedom.

With $\overline{\mathsf{P}}_1 = [I_3, \mathbf{0}]$ and the fact that the trifocal tensor (relation (7.7)) must be invariant of H, it is easy to see from (7.2) and (7.4) that the other two projection matrices must be given by

$$\begin{aligned} \overline{\mathsf{P}}_2 &\sim [\mathsf{A}, \mathbf{e}''_1]\,, \\ \overline{\mathsf{P}}_3 &\sim [\mathsf{B}, \mathbf{e}'''_1]\,. \end{aligned} \tag{7.41}$$

From these considerations we see that we can find directly from the defining quantities of the trifocal tensor a consistent set of projection matrices of the form (7.39) and (7.41), which are related to the real Euclidian configuration by some unknown projective transformation.

The epipoles $\mathbf{e}''_1$ and $\mathbf{e}'''_1$ can be computed following section 7.4.1, and the matrices A and B could then be determined linearly from the tensor equation (7.7); i.e. $T^{kj}_{\cdot\cdot i} \sim e''^{\,j}_1 B^k_{\cdot i} - e'''^{\,k}_1$. However, the solution for these two matrices is not unique. In the expression $T^{kj}_{\cdot\cdot i} \sim e''^{\,j}_1 B^k_{\cdot i} - e'''^{\,k}_1$

this can be seen by substituting A with $\mathsf{A} + \mathbf{e}_1''\mathbf{v}^\top$, B with $\mathsf{B} + \mathbf{e}_1'''\mathbf{v}^\top$, $\mathbf{e}_1''$ with $w\mathbf{e}_1''$, and $\mathbf{e}_1'''$ with $w\mathbf{e}_1'''$, for any $\mathbf{v}$ and any $w \neq 0$. The two additional $\mathbf{e}_1''\mathbf{v}^\top$ and $\mathbf{e}_1'''\mathbf{v}^\top$ terms simply chancel out, and w only changes the scale of the tensor. This is the same four way ambiguity that we have in relation (7.39) for determining H; cf. relation (7.40).

So we can choose any particular solution in this 4-parameter family for the matrices A and B. Hartley proposes the following solution, which extracts A and B directly from the correlation slices l_i; [Hartley 1997]:

$$\begin{aligned} \mathsf{A} &= (\mathbf{e}_1''\mathbf{e}_1''^\top - I)\left[\mathsf{l}_1^\top\mathbf{e}_1''', \mathsf{l}_2^\top\mathbf{e}_1''', \mathsf{l}_3^\top\mathbf{e}_1'''\right], \\ \mathsf{B} &= \left[\mathsf{l}_1\mathbf{e}_1'', \mathsf{l}_2\mathbf{e}_1'', \mathsf{l}_3\mathbf{e}_1''\right], \end{aligned} \tag{7.42}$$

with $|\mathbf{e}_1''| = |\mathbf{e}_1'''| = 1$. Note, that in this particular solution the ambiguity is solved by demanding that the columns of A are perpendicular to the epipole $\mathbf{e}_1''$. Therefore, A is singular and in this special projective frame the second projection center lies in the plane at infinity.

7.4.4 Retrieving the common interior orientation of the three images

We saw in section 7.4.3, that the projection matrices can not be retrieved uniquely from the trifocal tensor. Consequently an *unknown* interior orientation C_1, C_2 and C_3 can not be retrieved uniquely either. If, however, the interior orientation of the three images is known to be the same for all three images, then it can be retrieved in general. One way of doing this will be summarized in the following. This method is based on the so-called *Kruppa equations*, named after an Austrian mathematician and first used in [Faugeras 1992] for camera calibration. But prior to explaining these equations, we need to describe a unique feature in each image – the so-called image of the absolute conic.

The *absolute conic* lies in the plane at infinity and it is made up of the points $\mathbf{X} \sim (P, Q, R, S, T)^\top$ for which holds: $T = 0$ and $U^2 + V^2 + W^2 = 0$. Therefore it only contains imaginary points. Its map in an image is called *image of the absolute conic*, represented by the symmetric matrix W. Due to the definition of the absolute conic it is easy to see that its map W in any image depends only on the interior orientation C of the respective image

$$\mathsf{W} \sim \mathsf{C}^\top\mathsf{C}.$$

Of more interest is the *dual image of the absolute conic*, represented by W^*

$$\mathsf{W}^* \sim \mathsf{C}^{-1}\mathsf{C}^{-\top}. \tag{7.43}$$

Observe, that C is defined by 5 elements and W^* is a homogenous symmetric 3×3 matrix, which also has 5 independent elements. This one-to-one relation between the elements of W^* and the 5 elements of the interior orientation in C is the reason why the dual image of the absolute conic is so important for camera calibration.

Now back to the Kruppa equations. These equations refer to an *image pair* $\{\Psi_1, \Psi_2\}$ and express the constraint that, epipolar lines, which are tangent to a conic in image Ψ_1, are mapped by any homography in $\mathbb{P}^2$ to epipolar lines in image Ψ_2, which are also tangent to the corresponding conic in image Ψ_2. If we especially consider the conic W^*, the dual image of the absolute conic, then we can represent this constraint as; e.g. [Hartley and Zisserman 2001]

$$\mathsf{S}(\mathbf{e}_1'')\mathsf{W}^{*\prime\prime}\mathsf{S}(\mathbf{e}_1'') \sim \mathsf{F}_{12}^\top\mathsf{W}^{*\prime}\mathsf{F}_{12}.$$

Although this constraint involves nine relations, only two are independent, therefore in [Hartley and Zisserman 2001] a shorter version of the Kruppa equations is presented:

$$\begin{pmatrix} \boldsymbol{u}_2^\top \mathsf{W}^{*\prime} \boldsymbol{u}_2 \\ -\boldsymbol{u}_1^\top \mathsf{W}^{*\prime} \boldsymbol{u}_2 \\ \boldsymbol{u}_1^\top \mathsf{W}^{*\prime} \boldsymbol{u}_1 \end{pmatrix} \times \begin{pmatrix} s_1^2 \boldsymbol{v}_1^\top \mathsf{W}^{*\prime\prime} \boldsymbol{v}_1 \\ s_1 s_2 \boldsymbol{v}_1^\top \mathsf{W}^{*\prime\prime} \boldsymbol{v}_2 \\ s_2^2 \boldsymbol{v}_2^\top \mathsf{W}^{*\prime\prime} \boldsymbol{v}_2 \end{pmatrix} = \mathbf{0},$$

where $\boldsymbol{u}_i$, $\boldsymbol{v}_i$ and s_i are the columns and singular values of the SVD of the fundamental matrix F_{12}. It is easy to show that these three relations can also be written as

$$\begin{aligned} \boldsymbol{u}_1^\top \mathsf{W}^{*\prime} \mathsf{F}_{12} \mathsf{W}^{*\prime\prime} \boldsymbol{v}_2 &= 0, \\ \boldsymbol{u}_2^\top \mathsf{W}^{*\prime} \mathsf{F}_{12} \mathsf{W}^{*\prime\prime} \boldsymbol{v}_1 &= 0, \\ s_1 \boldsymbol{u}_1^\top \mathsf{W}^{*\prime} \mathsf{F}_{12} \mathsf{W}^{*\prime\prime} \boldsymbol{v}_1 - s_2 \boldsymbol{u}_2^\top \mathsf{W}^{*\prime} \mathsf{F}_{12} \mathsf{W}^{*\prime\prime} \boldsymbol{v}_2 &= 0. \end{aligned} \tag{7.44}$$

So we see, using the Kruppa equations we get two independent non-linear constraints on the elements of the dual images of the absolute conic $\mathsf{W}^{*\prime}$ and $\mathsf{W}^{*\prime\prime}$.

If the interior orientation of the three images associated with the trifocal tensor are unknown but common i.e. $\mathsf{C} = \mathsf{C}_1 = \mathsf{C}_2 = \mathsf{C}_3$, then we have 5 unknown elements in C. From the trifocal tensor we can easily obtain three image pairs with their respective fundamental matrices F_{12}, F_{13} and F_{23}; see section 7.4.2. Due to the common interior orientation also the dual images of the absolute conic are the same $\mathsf{W}^* = \mathsf{W}^{*\prime} = \mathsf{W}^{*\prime\prime} = \mathsf{W}^{*\prime\prime\prime}$ and by applying the Kruppa equations (7.44) onto the three image pairs, we get 9 equations in the 5 unknown elements of W^*, and 6 of these equations are independent. After determining W^* the one-to-one relation of W^* and C can be used to compute the common interior orientation of the three images by Cholesky factorization of W^*.

Since we have 9 equations in 5 unknowns we can solve the Kruppa equations in a least squares manner, provided initial values for the common interior orientation C are available (this then initializes W^* using (7.43))

$$M \Delta \boldsymbol{w} + \boldsymbol{m} = \boldsymbol{d} \quad \text{with} \quad |\boldsymbol{d}|^2 \rightarrow \min, \tag{7.45}$$

M is the Jacobian of the Kruppa equations with respect to the essential elements of W^* which make up the vector $\boldsymbol{w}$. The latter is iteratively refined by $\Delta \boldsymbol{w}$. $\boldsymbol{m}$ is the left side of the 6 Kruppa equations evaluated with the actual value of W^* and the fixed fundamental matrices F_{12}, F_{13} and F_{23}. The scale of W^* is either fixed by setting its Frobenius norm to 1 (and thereby adding an additional scaling constraint into system (7.45)) or by setting the element at position (3,3) to 1. The latter is a safe choice, because this element is inverse proportional to the square of the principal distance.

If the trifocal tensor itself was derived from noisy image correspondences (see section 8), then the interior orientation C determined in this way is not an optimal solution, because it is obtained by minimizing *algebraic error* and not by minimizing the errors in the original image observations (the so-called *reprojection error*). If an optimal solution of the interior orientation is the prime goal, then a subsequent bundle adjustment should be performed, where further image distortion parameters can be added. The values for C obtained by the presented method, however, are in general very good approximations for initializing this bundle adjustment.

Note, however, that for certain configurations of the three images not all 5 parameters of C can be determined. So, if there is no rotation between the images then this method will fail and if all images have the same rotation axis, then not all 5 parameters can be determined.

One open question is how to get the approximations of the interior orientation to initialize the iterative solution of system (7.45). The affine terms α and β can be assumed safely with 1 and 0 respectively. The principal point is usually near the center of the image[4]. So the approximations for these four elements should not be a problem in general. The initial value for the principal distance can be a bit difficult. Often we know the focal length used while taking the images, this will be a very good approximation for the principal distance. If the images are from an unknown source and there is no information on the principal distance or focal length, then using the approximations mentioned for α, β and the principal point, the Kruppa equations can be set up for an image pair with the latter interior elements fixed. The Kruppa equations then turn out to be *linear in the unknown principal distance* and so an approximation can be obtained quite easily. For more information on this topic of *self calibration* see again [Hartley and Zisserman 2001].

7.4.5 Retrieving the relative orientation of three calibrated images

If the interior orientation is known, either it is given for all images, or the common unknown interior orientation is determined following section 7.4.4, then it is straight forward to determine the quantities of the relative orientation of dependent images with respect to image Ψ_1, see definition 6.2 on page 45; i.e. $\mathbf{Z}_2{}^\dagger$, $R_2{}^\dagger$, $\mathbf{Z}_3{}^\dagger$, and $R_3{}^\dagger$.

Using the methods of section 7.4.2 we can retrieve the fundamental matrices F_{12} and F_{13} from the trifocal tensor. Together with the known interior orientation C_1, C_2 and C_3, and relation 6.12 we also get the respective essential matrices E_{12} and E_{13}. From these two matrices, one corresponding point in each image and the methods of section 6.1.4 on page 50 we can derive the rotation matrices $R_2{}^\dagger$ and $R_3{}^\dagger$ uniquely. Also the projection centers $\mathbf{Z}_2{}^\dagger$ and $\mathbf{Z}_3{}^\dagger$ are obtained with the correct sign, but with wrong relative scale as the length of *both* vectors is set to 1.

Using one pair of corresponding points $\mathbf{x}''$ and $\mathbf{x}'''$ in image Ψ_2 and Ψ_3, we can adapt the length of $\mathbf{Z}_3{}^\dagger$ to the unit vector $\mathbf{Z}_2{}^\dagger$ by a scale m. Starting from the relation

$$\mathbf{Z}_2{}^\dagger + \mu_2 R_2{}^\dagger C_2 \mathbf{x}'' = m\mathbf{Z}_3{}^\dagger + \mu_3 R_3{}^\dagger C_3 \mathbf{x}''',$$

we obtain the scale m as

$$m = \frac{|\mathbf{x}'''^\top C_3^\top R_3{}^{\dagger\top} \mathsf{S}(R_2{}^\dagger C_2 \mathbf{x}'')\mathbf{Z}_2{}^\dagger|}{|\mathbf{x}'''^\top C_3^\top R_3{}^{\dagger\top} \mathsf{S}(R_2{}^\dagger C_2 \mathbf{x}'')\mathbf{Z}_3{}^\dagger|}, \tag{7.46}$$

and the correctly scaled third projection center as $m\mathbf{Z}_3{}^\dagger$.

[4]It is important to point out that if all images are supposed to have the same interior orientation, then the measurement system (see section 5.1) must be the same in all images. This is true if the images are acquired with the same digital camera. However, if the images are shot using analog cameras and the measurement system refers to the analog images or to digital images obtained by scanning the analog ones, then all these images must be transformed into a common image coordinate system prior to computing the trifocal tensor. This common coordinate system can be defined by fiducial marks, which are part of professional cameras, or by the corners of the image frame (if visible) or by other striking points of the camera frame visible in all images.

7.4.6 Transformation of the individual relative orientations into one common frame in case of $n > 3$ images

Assume we are given $n > 3$ images together with corresponding image points, which are the mappings of some object points $\mathbf{X}_j$. Then we can compute $m < n$ trifocal tensors for various image triples with the methods of section 8 on page 111, so that each image is part of at least one tensor. Afterwards, following section 7.4.5, we can retrieve the parameters of the relative orientation, i.e. the basis vectors and rotation matrices, for each image triple with respect to the *first* image of each triple, provided the interior orientation is known or can be retrieved following section 7.4.4 on page 91.

Before all n images can be used simultaneously to perform spatial reasoning, it is required to transform all these individual relative orientations into one common system. This can be done basically in two ways; for simplicity we assume that we have $n = 9$ images $\Psi_1, \ldots, \Psi_9$:

(i) The m trifocal tensors are *not disjoint*; e.g. we have computed the following $m = 4$ tensors and obtained the respective parameters of the relative orientation:

$$\begin{aligned}
T(\Psi_1, \Psi_2, \Psi_3) &\rightarrow \mathbf{Z}_i^{[1]}, R_i^{[1]} \quad i = 1, 2, 3, \\
T(\Psi_3, \Psi_4, \Psi_5) &\rightarrow \mathbf{Z}_i^{[2]}, R_i^{[2]} \quad i = 3, 4, 5, \\
T(\Psi_5, \Psi_6, \Psi_7) &\rightarrow \mathbf{Z}_i^{[3]}, R_i^{[3]} \quad i = 5, 6, 7, \\
T(\Psi_7, \Psi_8, \Psi_9) &\rightarrow \mathbf{Z}_i^{[4]}, R_i^{[4]} \quad i = 7, 8, 9.
\end{aligned}$$

The super-script $[f]$ indicates the frame, defined by the first image of the fth tensor, in which the respective entity is given. In this case all orientation parameters can be transformed into the frame of the first tensor in the following way:

$$\begin{aligned}
\mathbf{Z}_i^{[1]} &= R_a^{[1]} R_a^{[f]\top} (\mathbf{Z}_i^{[f]} - \mathbf{Z}_a^{[f]}) + \mathbf{Z}_a^{[1]} \\
R_i^{[1]} &= R_a^{[1]} R_a^{[f]\top} R_i^{[f]},
\end{aligned}$$

where $f = 2, 3, 4$ and a is the number of the image common to frame $[1]$ and frame $[f]$. We therefore have $a = 3$ for the transformation of frame $[2]$, $a = 5$ for frame $[3]$, and $a = 7$ for frame $[4]$.

(ii) The m trifocal tensors are *disjoint*. In this case the basis vectors and rotation matrices of the relative orientation can not be used directly to compute the transformation from each frame into the frame of the first tensor. However, if the sets of points used to compute the individual tensors are not disjoint, then the transformation can be derived from the respective object points $\mathbf{X}_j$, which are obtained for each individual image triple by spatial intersection. For our example we have computed the following $m = 3$ tensors and derived the respective object points, basis vectors and rotation matrices:

$$\begin{aligned}
T(\Psi_1, \Psi_2, \Psi_3) &\rightarrow \mathbf{X}_{j'}^{[1]} \text{ and } \mathbf{Z}_i^{[1]}, R_i^{[1]}, \quad i = 1, 2, 3, \\
T(\Psi_4, \Psi_5, \Psi_6) &\rightarrow \mathbf{X}_{j''}^{[2]} \text{ and } \mathbf{Z}_i^{[2]}, R_i^{[2]}, \quad i = 4, 5, 6, \\
T(\Psi_7, \Psi_8, \Psi_9) &\rightarrow \mathbf{X}_{j'''}^{[3]} \text{ and } \mathbf{Z}_i^{[3]}, R_i^{[3]}, \quad i = 7, 8, 9.
\end{aligned}$$

Now we compute the transformation parameters $R_{[f]}^{[1]}$ and $Q_{[1]}^{[f]}$, which take the object points of each individual frame into the frame of the first tensor as $\mathbf{X}_j^{[1]} = R_{[f]}^{[1]}(\mathbf{X}_j^{[f]} - Q_{[1]}^{[f]})$, with $f = 2, 3$. This transformation is uniquely determined if at least 3 non-collinear points, common to frame [1] and frame [f], are given.

The basis vectors and rotation matrices are then transformed as:

$$\begin{aligned} \mathbf{Z}_i^{[1]} &= R_{[f]}^{[1]}(\mathbf{Z}_i^{[f]} - Q_{[1]}^{[f]}) \\ R_i^{[1]} &= R_{[f]}^{[1]} R_i^{[f]}. \end{aligned}$$

The first method, which directly uses the basis vectors and rotation matrices, is easier to handle than the second one, because it does not need any object points. However, for applying the first method more trifocal tensors for the same number of images must be computed than for the second method. If n images are given, then for the first method at least $\lceil \frac{n-1}{2} \rceil$ trifocal tensors and for the second method at least $\lceil \frac{n}{3} \rceil$ trifocal tensors must be computed.

7.5 Changing the order of the images

In section 7.1 we saw that there are actually six different trifocal tensors associated with three given views Ψ_1, Ψ_2 and Ψ_3. If we have computed the trifocal tensor $T(\Psi_1, \Psi_2, \Psi_3)$ then the tensors for the five other image orders can be derived directly from the computed one. In [Avidan and Shashua 1996] this task is accomplished by introducing two operators: O_{12}, which exchanges the first and the second view, and O_{23}, which exchanges the second and the third view.

The latter operator O_{23} is simply realized. The exchange of the second and third view results in interchanging the J and K homographies. This can be done just by transposing the I matrices. Thus $T(\Psi_1, \Psi_3, \Psi_2) = \{\overline{\mathsf{I}_1}, \overline{\mathsf{I}_2}, \overline{\mathsf{I}_3}\}$ is obtained from $T(\Psi_1, \Psi_2, \Psi_3) = \{\mathsf{I}_1, \mathsf{I}_2, \mathsf{I}_3\}$ using the operator O_{23} as $T(\Psi_1, \Psi_3, \Psi_2) = \{\mathsf{I}_1{}^\top, \mathsf{I}_2{}^\top, \mathsf{I}_3{}^\top\}$.

The operator O_{12} is a bit more involving. The K homographies of the given tensor $T(\Psi_1, \Psi_2, \Psi_3) = \{\mathsf{K}_1, \mathsf{K}_2, \mathsf{K}_3\}$ map the points of image Ψ_1 to the points of image Ψ_2 via the principal planes of image Ψ_3. The $\overline{\mathsf{K}}$ homographies of the desired tensor $T(\Psi_2, \Psi_1, \Psi_3) = \{\overline{\mathsf{K}_1}, \overline{\mathsf{K}_2}, \overline{\mathsf{K}_3}\}$ map the points of image Ψ_2 to the points of image Ψ_1 via the same principal planes of image Ψ_3. This mapping corresponds to the inverse of the given K homographies. Thus it holds $\overline{\mathsf{K}_k} \sim \mathsf{K}_k{}^{-1}$, or $\overline{\mathsf{K}_k} = \mu_k \mathsf{K}_k{}^{-1}$. Of the three remaining unknown scales μ_k only two are essential, because one can be cancelled for the overall scaling of the tensor. In [Avidan and Shashua 1996] these two scales are determined using a measure called *projective depth* for a pair of corresponding points in the images Ψ_1 and Ψ_2. Note that this operator O_{12} can run into problems if the K homographies are singular. In [Heyden 2000] and [Faugeras and Luong 2001] basically the same approach is used. There, however, the adjoint matrices of the K homographies are considered, which allows to cope with rank-2 homographies; but still leaves an open problem for rank-1 homographies.

In [Hartley and Zisserman 2001] a totally different method is presented for deriving the trifocal tensor from a given one if the reference view is changed. Their method does not suffer from possible singular homographies, and is outlined as follows: (a) Extract the three projection

matrices P_1, P_2 and P_3 from the tensor, by the method of section 7.4.3; i.e. with $\mathsf{P}_1 = [I_3, \mathbf{0}]$. (b) Find a regular projective transformation H such that $\mathsf{P}_2\mathsf{H} = [I_3, \mathbf{0}]$; i.e. solve $(I_4 \otimes \mathsf{P}_2)\,\mathrm{vec}\,(\mathsf{H}) = \mathrm{vec}\,([I_3, \mathbf{0}])$. Since this system has only 12 equations for the 16 unknowns of H, the solution for H is a four-parameter family. Select any parameter-set that produces a regular H and apply H to P_1 and P_3. (c) Compute the trifocal tensor $T(\Psi_2, \Psi_1, \Psi_3)$ with these *transformed* projection matrices following section 7.1.1.

7.6 The internal constraints

We saw in section 7.1 during the derivation of the trifocal tensor, that it is made up of 27 elements but has only 18 degrees of freedom. Therefore 9 constraints have to be satisfied by these 27 elements to represent a valid trifocal tensor. One of these constraints is the fixing of the scale of the homogenous tensor (e.g. by setting its Frobenius norm to 1) and therefore 8 constraints are essentially to be considered. These constraints are especially important when the tensor is computed from given point and/or line correspondences across three views; see section 8.

Over the past few years a lot of research has gone into finding these constraints. Some of the presented constraints are not complete (<8), some are not minimal (>8) and others are rather complicated. In the following the most important sets of constraints presented in the literature so far are described. Afterwards two new sets are derived and from these we will see that the geometric meaning behind the internal constraints is rather simple: The internal constraints of the trifocal tensor just guarantee that the maps of the principal rays $\boldsymbol{\mathcal{R}}_i$ of each image into the other images are straight and concurrent lines.

All constraints are in some way related to the properties of the tensorial slices presented in section 7.2 and they are derived based on the assumption that the images are in general position – in this case the tensorial slices will have maximum rank.

7.6.1 The non-minimal set of Papadopoulo and Faugeras

Perhaps the first set of constraints that fully characterize the trifocal tensor was proposed by Papadopoulo and Faugeras in [Papadopoulo and Faugeras 1998]. It is a non-minimal set consisting of 12 dependent equations – the so-called *epipolar* and *extended rank constraints*, which are derived from the correlation slices l_i.

Proposition 7.14 (The rank and epipolar constraints.) *The correlation slices l_i of the trifocal tensor satisfy the three **rank constraints***

$$\det(\mathsf{l}_i) = 0. \tag{7.47}$$

*The trifocal tensor also satisfies the two **epipolar constraints***

$$\det(\mathsf{N}_R) = \det(\mathsf{N}_L) = 0, \tag{7.48}$$

where N_R and N_L are made up of the right and left nullspaces of the l_i; see proposition 7.4 on page 69.

These constraints, which have been presented also by many other authors e.g. [Spetsakis and Aloimonos 1990], [Hartley 1994b], or [Stein and Shashua 1998], follow directly from the properties of the correlation slices l_i in the propositions 7.3 and 7.4.

The epipolar constraints got their name from the fact that the right and left nullspaces $\mathcal{N}(\mathsf{l}_i)$ and $\mathcal{N}(\mathsf{l}_i^\top)$, provided $\operatorname{rank}(\mathsf{l}_i) = 2$, represent the lines $\boldsymbol{\rho}''_{1i}$ and $\boldsymbol{\rho}'''_{1i}$ respectively. These lines are the mappings of the principal rays of image Ψ_1 into the images Ψ_2 and Ψ_3 respectively. Therefore these lines are epipolar lines and consequently they must be concurrent through the respective epipole. It is, however, important to point out, that these epipolar constraints are only applicable if all three correlation slices l_i have maximum rank-2, because only in this case the nullspaces are lines. If $\operatorname{rank}(\mathsf{l}_i) < 2$ the nullspaces are either a pencil of lines through the respective epipole or the entire $\mathbb{P}^2$; cf. table 7.1. From table 7.4 on page 77 we see that $\operatorname{rank}(\mathsf{l}_i) < 2$ can happen for any image configuration. However, since these ranks depend on the incidence of the projection centers $\mathbf{Z}_2$ and $\mathbf{Z}_3$ with the principal rays of image Ψ_1 (see proposition 7.3 on page 68), by proper choice of the measurement system of image Ψ_1 the general case of rank-2 for all three correlation slices can be achieved generally. Although it may not be obvious in advance how to choose the coordinate system in image Ψ_1.

These five rank and epipolar constraints have a simple geometric interpretation, which was given at the end of section 7.2.1 and are independent, since the singularity of the matrices says nothing about the way the nullspaces are related. However, they are not sufficient to describe the internal constraints of the trifocal tensor because for that eight constraints are required. In [Papadopoulo and Faugeras 1998], however, the authors also introduce the so-called *extended rank constraints*.

Proposition 7.15 (The extended rank constraints.) *The correlation slices l_i of the trifocal tensor satisfy the ten* ***extended rank constraints***

$$\operatorname{rank}\Big(\sum_{i=1}^{3} x_i \mathsf{l}_i\Big) \leq 2, \quad \textit{for all } x_i, i = 1, 2, 3. \tag{7.49}$$

These constraint correspond to proposition 7.2 on page 67, which says that the linear combination of the correlation slices l_i is also a singular dual correlation matrix – the one induced by the projection ray of the point $(x_1, x_2, x_3)^\top$ in image Ψ_1. Note that the three rank constraints of (7.47) are of course included in the extended ones. To see that (7.49) are ten constraints, we have to expand $\det(\sum_{i=1}^{3} x_i \mathsf{l}_i) = 0$ with respect to the individual products of the x_i. The resulting polynomial of degree three in the x_i has $\binom{3+3-1}{3} = 10$ individual terms. Since (7.49) must hold for any choice of the x_i, all ten coefficients (which are sums of three determinants made up of the columns of the l_i matrices) of this polynomial must vanish individually.

Papadopoulo and Faugeras proof that any homogenous valence-(1,2) tensor that satisfies the ten extended rank and the two epipolar constraints is a trifocal tensor. In other words these 12 constraints represent a non-minimal set of constraints that fully characterize the trifocal tensor. The extended rank constraints are of degree 3 in the elements of the tensor and the epipolar constraints are of degree 6.

7.6.2 The minimal set of Canterakis

The first minimal set of constraints was presented by Canterakis in [Canterakis 2000]. His 8 constraints are based on the properties of the homography slices J_j given in proposition 7.8

(and since the second and third image play similar roles with respect to the trifocal tensor, corresponding constraints could be derived from the homography slices K_k as well):

- The polynomial $\det(\mathsf{J}_2 - \kappa\mathsf{J}_1) = 0$ should have a double root $\kappa_1 = \kappa_2$ with $\mathsf{J}_2 - \kappa_1\mathsf{J}_1$ having rank-1, and a single root κ_3.
- The polynomial $\det(\mathsf{J}_3 - \overline{\kappa}\mathsf{J}_1) = 0$ should have a double root $\overline{\kappa}_1 = \overline{\kappa}_2$ with $\mathsf{J}_2 - \overline{\kappa}_1\mathsf{J}_1$ having rank-1, and a single root $\overline{\kappa}_3$.
- The general eigenvectors of the single roots κ_3 and $\overline{\kappa}_3$ should be the same.

The algebraic representation of these constraints is not straightforward and Canterakis proposes the following. The polynomial $\det(\mathsf{J}_2 - \kappa\mathsf{J}_1) = 0$ can be represented as $a\kappa^3 + b\kappa^2 + c\kappa + d = 0$, where a, b, c, d are expressions in the elements of the tensor (i.e. in J_1 and J_2). If this polynomial is supposed to have a double root, then using $A = b^2 - 3ac$, $B = bc - 9ad$ and $C = c^2 - 3bd$ the following condition has to be satisfied

$$B^2 - 4AC = 0, \tag{7.50}$$

which is of degree 12 in the elements of the tensor. Using the quantities A, B and C the roots can be expressed directly as

$$\kappa_1 = \kappa_2 = -\frac{B}{2A} \quad \text{and} \quad \kappa_3 = \frac{B}{A} - \frac{b}{a}.$$

However, no direct representation for the corresponding eigenvectors is given in [Canterakis 2000]. Let us denote the two dimensional eigenvector with $\boldsymbol{\gamma}$ (a line) and the single one with $\mathbf{c}$ (a point). We know from proposition 7.8 that $\boldsymbol{\gamma} \sim \boldsymbol{\rho}'_{23}$ (the map of the principal ray $\boldsymbol{\mathcal{R}}''_3$ of image Ψ_2 into image Ψ_1), and that $\mathbf{c} \sim \mathbf{e}'_3$ (the epipole in image Ψ_1 with respect to image Ψ_3).

The condition that $\mathsf{J}_2 - \kappa_1\mathsf{J}_1$ has rank-1 can be checked as follows. First let us recall that due to the homography-epipole-relation (7.16) on page 72 $\mathbf{c} \sim \mathbf{e}'_3$ is mapped via any regular J_j to $\mathbf{e}'''_1$. If the rank-1 condition is satisfied then we see from table 7.2 that $\mathsf{J}_2 - \kappa_1\mathsf{J}_1$ can be written as $\mathsf{J}_2 - \kappa_1\mathsf{J}_1 = \mathbf{e}'''_1\boldsymbol{\gamma}^\top$. Therefore any point $\mathbf{x} \notin \boldsymbol{\gamma}$ is mapped via $\mathsf{J}_2 - \kappa_1\mathsf{J}_1$ to $\mathbf{e}'''_1$. Because the point $\mathbf{x} = \boldsymbol{\gamma}^*$ (the dual of $\boldsymbol{\gamma}$) will not lie on $\boldsymbol{\gamma}$, the rank-1 condition can be tested by

$$(\mathsf{J}_2 - \kappa_1\mathsf{J}_1)\boldsymbol{\gamma} \sim \mathsf{J}_1\mathbf{c}, \tag{7.51}$$

which is equivalent to two equations.

Summing it up the 8 constraints of Canterakis are realized in the following way:

- Two constraints of the form (7.50) by demanding double roots for the polynomials $\det(\mathsf{J}_2 - \kappa\mathsf{J}_1) = 0$ and $\det(\mathsf{J}_3 - \overline{\kappa}\mathsf{J}_1) = 0$.
- Two constraints of the form (7.51) by demanding rank-1 for $\mathsf{J}_2 - \kappa_1\mathsf{J}_1$.
- Two constraints of the form (7.51) by demanding rank-1 for $\mathsf{J}_3 - \overline{\kappa}_1\mathsf{J}_1$.
- Two constraints by demanding that the eigenvectors of the single roots are the same: $\mathbf{c} \sim \overline{\mathbf{c}}$.

For applying this set of constraints we need at least one regular homography slice which plays the role of the matrix J_1 in the previous relations, and from table 7.4 on page 77 we see that this is possible as long as the second projection centers is distinct from the other two, which can be reached by adapting the roles of the given images properly.

7.6.3 The new minimal sets of constraints

In the following we will derive two new minimal sets, one from the correlation slices I_i and one from the homography slices J_j. From the homography slices K_k we can derive similar constraints to those obtained from J_j, just be interchanging the roles of the second and the third image – therefore we will only consider the J_j matrices. However, this just shows that minimal sets of constraints can be derived from any kind of slice. This is not very surprising after all, as all 27 elements of the tensor are involved in the three matrices of each kind of slice.

Both new sets are guided by the *same geometric consideration*, that the maps of the principal rays of each image into the other images have to be straight and concurrent lines, see figure 7.8, and we adopt the naming convention given in table 7.8.

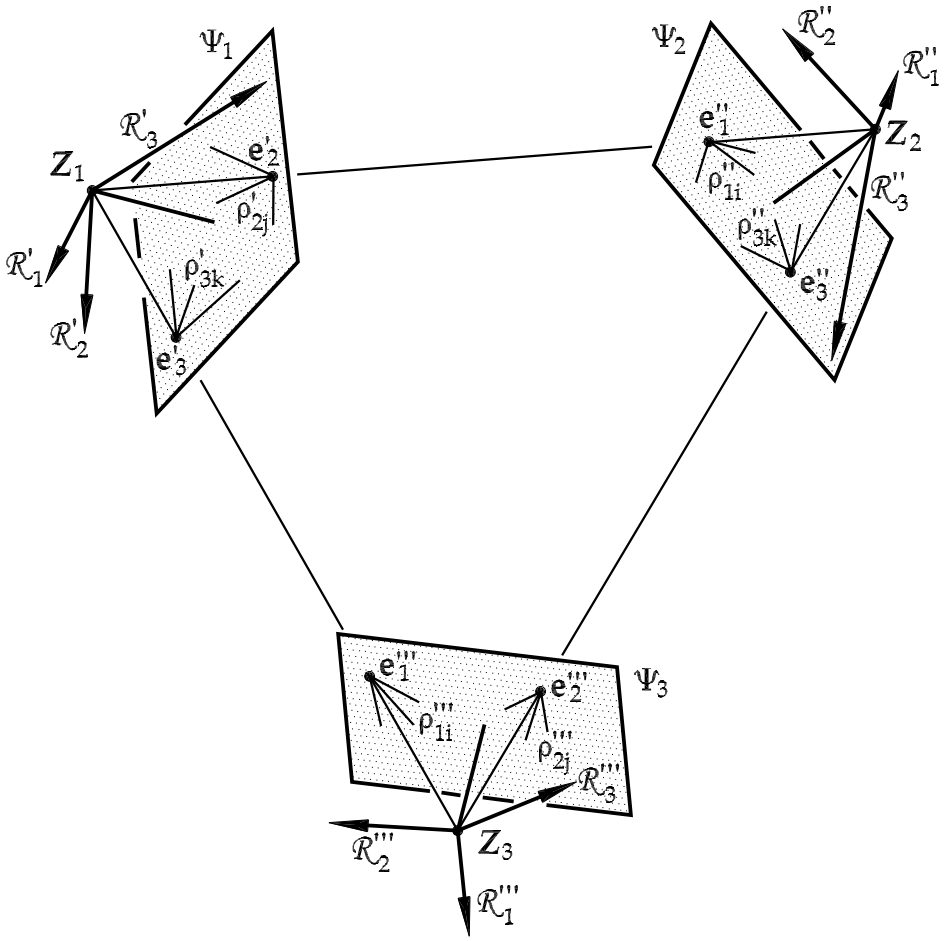

Figure 7.8: *The maps $\boldsymbol{\rho}_{\cdot\cdot}^{(\cdot)}$ of the principal rays $\boldsymbol{\mathcal{R}}_{\cdot}^{(\cdot)}$ of each image into the other two images.*

Although, this geometric consideration seems to be a rather trivial realization, it is exactly this condition, which is not satisfied, if the internal constraints of the tensor are neglected. We record this in the following theorem and prove it at the end of this section.

Theorem 7.1 (The geometric meaning of the internal constraints.) *The internal constraints of the trifocal tensor guarantee that the maps $\boldsymbol{\rho}_{\cdot\cdot}^{(\cdot)}$ of the principal rays $\boldsymbol{\mathcal{R}}_{\cdot}^{(\cdot)}$ of each image into the other images are straight and concurrent lines.*

$\boldsymbol{\rho}'_{2j}$	maps of the principal rays $\boldsymbol{\mathcal{R}}''_j$ of image Ψ_2 into image Ψ_1
$\boldsymbol{\rho}'_{3k}$	maps of the principal rays $\boldsymbol{\mathcal{R}}'''_k$ of image Ψ_3 into image Ψ_1
$\boldsymbol{\rho}''_{1i}$	maps of the principal rays $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1 into image Ψ_2
$\boldsymbol{\rho}''_{3k}$	maps of the principal rays $\boldsymbol{\mathcal{R}}'''_k$ of image Ψ_3 into image Ψ_2
$\boldsymbol{\rho}'''_{1i}$	maps of the principal rays $\boldsymbol{\mathcal{R}}'_i$ of image Ψ_1 into image Ψ_3
$\boldsymbol{\rho}'''_{2j}$	maps of the principal rays $\boldsymbol{\mathcal{R}}''_j$ of image Ψ_2 into image Ψ_3

Table 7.8: *The naming convention for the maps $\boldsymbol{\rho}^{(\cdot)}_{\cdot\cdot}$ of the principal rays $\boldsymbol{\mathcal{R}}^{(\cdot)}_{\cdot}$ of each image into the other two images.*

It is simple to transfer this geometric consideration into a minimal set of constraints, because the rows and columns of the correlation and homography slices correspond either to *points lying on* or to *lines intersecting on* the maps of the principal rays in the images. These points and lines are the mappings of 3D points or 3D lines which arise as the intersections of particular principal rays or principal planes of one image with the principal planes of another image. These relations are summarized in table 7.9.

slice	row-indx	col-indx	columns are	of ... in space	rows are	of ... in space
l_i	k	j	points in Ψ_3	$\boldsymbol{\mathcal{R}}'_i \wedge \boldsymbol{\Pi}''_j$	points in Ψ_2	$\boldsymbol{\mathcal{R}}'_i \wedge \boldsymbol{\Pi}'''_k$
J_j	k	i	points in Ψ_3	$\boldsymbol{\mathcal{R}}'_i \wedge \boldsymbol{\Pi}''_j$	lines in Ψ_1	$\boldsymbol{\Pi}''_j \wedge \boldsymbol{\Pi}'''_k$
K_k	j	i	points in Ψ_2	$\boldsymbol{\mathcal{R}}'_i \wedge \boldsymbol{\Pi}'''_k$	lines in Ψ_1	$\boldsymbol{\Pi}''_j \wedge \boldsymbol{\Pi}'''_k$

Table 7.9: *The geometric interpretation of the rows and columns of the three tensorial slices as mapped points and lines in the images Ψ_1, Ψ_2 and Ψ_3. The principal rays of image Ψ_1 are given as $\boldsymbol{\mathcal{R}}'_i$. The principal planes of the images Ψ_2 and Ψ_3 are given as $\boldsymbol{\Pi}''_j$ and $\boldsymbol{\Pi}'''_k$ respectively.*

For simpler handling we will refer to the columns and rows of the three slices individually by assigning the names given in table 7.10 on the facing page.

A new minimal set derived from the correlation slices

Let us begin with the correlation slices l_i. From table 7.9 we see that the jth column of l_i is a point in image Ψ_3, which is the mapping of the intersection of the principal plane $\boldsymbol{\Pi}''_j$ with the principal ray $\boldsymbol{\mathcal{R}}'_i$. Consequently all three columns of each l_i matrix must be collinear – lying on the epipolar line $\boldsymbol{\rho}'''_{1i}$, which is the map of the the principal ray $\boldsymbol{\mathcal{R}}'_i$ into image Ψ_3, see figure 7.10 on page 104 left part. The algebraic expression for these collinear properties is that the determinants of the correlation matrices must vanish

$$\begin{aligned} |\mathbf{a}'''_1, \quad \mathbf{a}'''_2, \quad \mathbf{a}'''_3| &= \det(\mathsf{l}_1) = 0 \\ |\mathbf{b}'''_1, \quad \mathbf{b}'''_2, \quad \mathbf{b}'''_3| &= \det(\mathsf{l}_2) = 0 \\ |\mathbf{c}'''_1, \quad \mathbf{c}'''_2, \quad \mathbf{c}'''_3| &= \det(\mathsf{l}_3) = 0\,. \end{aligned} \tag{7.52}$$

$$
\begin{array}{ll}
\mathsf{I}_1 = \left[\, \mathbf{a}_1''', \;\; \mathbf{a}_2''', \;\; \mathbf{a}_3''' \,\right] & \mathsf{J}_1 = \left[\, \mathbf{a}_1''', \;\; \mathbf{b}_1''', \;\; \mathbf{c}_1''' \,\right] \\
\mathsf{I}_2 = \left[\, \mathbf{b}_1''', \;\; \mathbf{b}_2''', \;\; \mathbf{b}_3''' \,\right] & \mathsf{J}_2 = \left[\, \mathbf{a}_2''', \;\; \mathbf{b}_2''', \;\; \mathbf{c}_2''' \,\right] \\
\mathsf{I}_3 = \left[\, \mathbf{c}_1''', \;\; \mathbf{c}_2''', \;\; \mathbf{c}_3''' \,\right] & \mathsf{J}_3 = \left[\, \mathbf{a}_3''', \;\; \mathbf{b}_3''', \;\; \mathbf{c}_3''' \,\right] \\
\\
\mathsf{K}_1 = \left[\, \mathbf{d}_1'', \;\; \mathbf{e}_1'', \;\; \mathbf{f}_1'' \,\right] & \mathsf{I}_1^\top = \left[\, \mathbf{d}_1'', \;\; \mathbf{d}_2'', \;\; \mathbf{d}_3'' \,\right] \\
\mathsf{K}_2 = \left[\, \mathbf{d}_2'', \;\; \mathbf{e}_2'', \;\; \mathbf{f}_2'' \,\right] & \mathsf{I}_2^\top = \left[\, \mathbf{e}_1'', \;\; \mathbf{e}_2'', \;\; \mathbf{e}_3'' \,\right] \\
\mathsf{K}_3 = \left[\, \mathbf{d}_3'', \;\; \mathbf{e}_3'', \;\; \mathbf{f}_3'' \,\right] & \mathsf{I}_3^\top = \left[\, \mathbf{f}_1'', \;\; \mathbf{f}_2'', \;\; \mathbf{f}_3'' \,\right] \\
\\
\mathsf{J}_1^\top = \left[\, \boldsymbol{\alpha}_1', \;\; \boldsymbol{\alpha}_2', \;\; \boldsymbol{\alpha}_3' \,\right] & \mathsf{K}_1^\top = \left[\, \boldsymbol{\alpha}_1', \;\; \boldsymbol{\beta}_1', \;\; \boldsymbol{\gamma}_1' \,\right] \\
\mathsf{J}_2^\top = \left[\, \boldsymbol{\beta}_1', \;\; \boldsymbol{\beta}_2', \;\; \boldsymbol{\beta}_3' \,\right] & \mathsf{K}_2^\top = \left[\, \boldsymbol{\alpha}_2', \;\; \boldsymbol{\beta}_2', \;\; \boldsymbol{\gamma}_2' \,\right] \\
\mathsf{J}_3^\top = \left[\, \boldsymbol{\gamma}_1', \;\; \boldsymbol{\gamma}_2', \;\; \boldsymbol{\gamma}_3' \,\right] & \mathsf{K}_3^\top = \left[\, \boldsymbol{\alpha}_3', \;\; \boldsymbol{\beta}_3', \;\; \boldsymbol{\gamma}_3' \,\right]
\end{array}
$$

Table 7.10: *Assigning names to the columns and rows of the three tensorial slices, which are either points or lines in the images Ψ_1, Ψ_2 and Ψ_3. See table 7.9 for the interpretation of these columns and rows, and figure 7.9 for a visual representation.*

These three constraints are the rank constraints (7.47) on page 96 and just correspond to the property that the I_i matrices have at most rank-2 and are therefore always singular. Since the principal rays $\boldsymbol{\mathcal{R}}_i'$ are concurrent also their mappings – the epipolar line $\boldsymbol{\rho}_{1i}'''$ – have to be so. This gives another constraint – one of the epipolar constraints in (7.48) on page 96

$$
\left|\, \mathbf{a}_1''' \times \mathbf{a}_2''', \quad \mathbf{b}_1''' \times \mathbf{b}_2''', \quad \mathbf{c}_1''' \times \mathbf{c}_2''' \,\right| = 0. \tag{7.53}
$$

This constraint was set up with the first two columns of the I_i matrices, but in general can be set up with any pair of columns. Since these four constraints are identical to the rank and (half of the) epipolar constraints in section 7.6.1, we know they are independent.

So we need to find four more constraints to form a minimal set of constraints. These remaining constraints are new and use the property (c) in proposition 7.4 on page 69. If we adapt this property onto the transposed correlation slices $\mathsf{I}_i^\top$ we get the following lemma.

Lemma 7.1 *Any line $\boldsymbol{\lambda}''' \neq \boldsymbol{\rho}_{1i}'''$ through the epipole $\mathbf{e}_1'''$ in image Ψ_3 is mapped by $\mathsf{I}_i^\top$ to the epipole $\mathbf{e}_1''$ in image Ψ_2.*

With this property we can formulate the following relation:

$$
\mathsf{I}_1^\top \boldsymbol{\lambda}_1''' \sim \mathsf{I}_2^\top \boldsymbol{\lambda}_2''' \sim \mathsf{I}_3^\top \boldsymbol{\lambda}_3''', \tag{7.54}
$$

with arbitrary $\boldsymbol{\lambda}_i''' \neq \boldsymbol{\rho}_{1i}'''$ and $\mathbf{e}_1''' \in \boldsymbol{\lambda}_i'''$. We can represent lines $\boldsymbol{\lambda}_i'''$ going through $\mathbf{e}_1'''$ as $\boldsymbol{\lambda}_i''' \sim \mathsf{S}(\mathbf{e}_1''')\mathbf{p}_i'''$, with arbitrary $\mathbf{p}_i''' \notin \boldsymbol{\rho}_{1i}'''$. And a point $\mathbf{p}_i'''$ that satisfies the condition $\mathbf{p}_i''' \notin \boldsymbol{\rho}_{1i}'''$ is simply the dual of $\boldsymbol{\rho}_{1i}'''$; i.e. $\mathbf{p}_i''' = \boldsymbol{\rho}_{1i}'''^*$. So we can rewrite the relations (7.54) and get the following **mapping constraints**

$$
\mathsf{I}_1^\top \mathsf{S}(\mathbf{e}_1''')\boldsymbol{\rho}_{11}''' \sim \mathsf{I}_2^\top \mathsf{S}(\mathbf{e}_1''')\boldsymbol{\rho}_{12}''' \sim \mathsf{I}_3^\top \mathsf{S}(\mathbf{e}_1''')\boldsymbol{\rho}_{13}'''. \tag{7.55}
$$

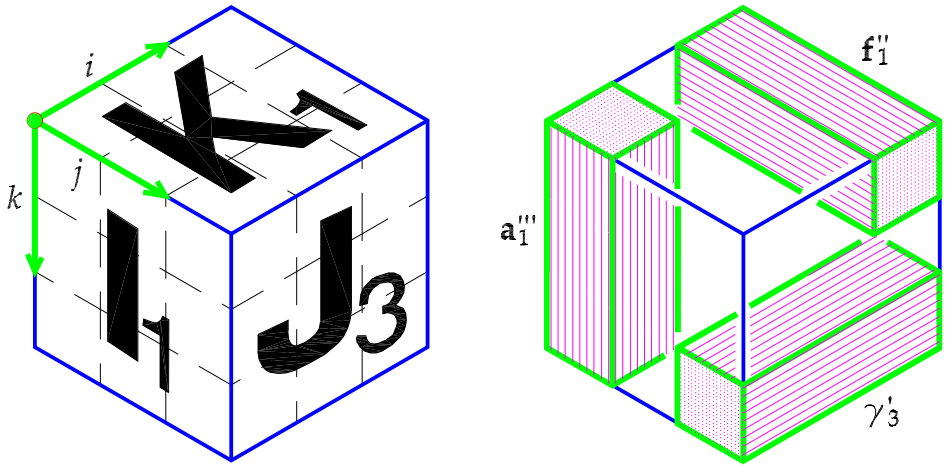

Figure 7.9: *A visual representation of the vectors $\mathbf{a}_1'''$, $\mathbf{f}_1''$ and $\boldsymbol{\gamma}_3'$. The vector $\mathbf{a}_1'''$ is the first column of I_1 and the first column of J_1. The vector $\mathbf{f}_1''$ is the first row of I_3 and the last column of K_1. The vector $\boldsymbol{\gamma}_3'$ is the last row of J_3 and the last row of K_3.*

Note that the additional entities in this relation, i.e. the epipole $\mathbf{e}_1'''$ and the three epipolar lines $\boldsymbol{\rho}_{1i}'''$, can be expressed by cross products of the columns of the I_i matrices. The three $\boldsymbol{\rho}_{1i}'''$ are the columns of the matrix in the epipolar constraint (7.53) and the epipole $\mathbf{e}_1'''$ is the cross product of two convenient $\boldsymbol{\rho}_{1i}'''$; e.g.

$$\mathbf{e}_1''' \sim (\mathbf{a}_1''' \times \mathbf{a}_2''') \times (\mathbf{b}_1''' \times \mathbf{b}_2''').$$

The relation (7.55) corresponds to six relations with two unknown scales μ and $\overline{\mu}$, and could be reformulated for example as

$$\begin{aligned}\mu \mathsf{I}_1^\top \mathsf{S}(\mathbf{e}_1''')\boldsymbol{\rho}_{11}''' &= \mathsf{I}_2^\top \mathsf{S}(\mathbf{e}_1''')\boldsymbol{\rho}_{12}''' \\ \overline{\mu} \mathsf{I}_1^\top \mathsf{S}(\mathbf{e}_1''')\boldsymbol{\rho}_{11}''' &= \mathsf{I}_3^\top \mathsf{S}(\mathbf{e}_1''')\boldsymbol{\rho}_{13}''',\end{aligned}$$

giving four constraints in total. Since each of these relations involves two different I_i matrices, it is obvious that they are independent. What remains to show is that all eight constraints (7.52), (7.53) and (7.55) are independent among each other. This will be shown next.

Assume we have three matrices I_i, which are supposed to be correlation slices of a trifocal tensor, and we know that these matrices satisfy the four independent constraints (7.52) and (7.53), then we can parameterize them in the following way without loss of generality

$$\mathsf{I}_i = \left[\mathbf{s}_i''', \quad v_i\mathbf{s}_i''' + m_i\mathbf{e}_1''', \quad w_i\mathbf{s}_i''' + n_i\mathbf{e}_1'''\right]. \tag{7.56}$$

This parameterization ensures that the columns of each matrix I_i are collinear, because they are spanned by two vectors $\mathbf{s}_i'''$ and $\mathbf{e}_1'''$, and that the carrier lines (i.e. the epipolar lines $\boldsymbol{\rho}_{1i}'''$) of these columns are concurrent in $\mathbf{e}_1'''$. Now we have to check if these matrices satisfy lemma 7.1. If the lemma is satisfied, then the constraints (7.55) are dependent on the first four constraints. If it is not satisfied, then all eight constraints are independent. Therefore we multiply each matrix $\mathsf{I}_i^\top$ with a line $\boldsymbol{\lambda}_i'''$ that is different from the carrier lines of the I_i matrices ($\rightarrow \mathbf{s}_i'''^\top \boldsymbol{\lambda}_i''' \neq 0$) and which goes through the common point $\mathbf{e}_1'''$. Taking into account that $\mathbf{e}_1'''^\top \boldsymbol{\lambda}_i''' = 0$ we therefore

get:

$$\mathsf{l}_1^\top \boldsymbol{\lambda}_1''' = \begin{pmatrix} \mathbf{s}_1'''^\top \boldsymbol{\lambda}_1''' \\ v_1 \mathbf{s}_1'''^\top \boldsymbol{\lambda}_1''' \\ w_1 \mathbf{s}_1'''^\top \boldsymbol{\lambda}_1''' \end{pmatrix} \sim \begin{pmatrix} 1 \\ v_1 \\ w_1 \end{pmatrix}$$
$$\mathsf{l}_2^\top \boldsymbol{\lambda}_2''' = \begin{pmatrix} \mathbf{s}_2'''^\top \boldsymbol{\lambda}_2''' \\ v_2 \mathbf{s}_2'''^\top \boldsymbol{\lambda}_2''' \\ w_2 \mathbf{s}_2'''^\top \boldsymbol{\lambda}_2''' \end{pmatrix} \sim \begin{pmatrix} 1 \\ v_2 \\ w_2 \end{pmatrix}$$
$$\mathsf{l}_3^\top \boldsymbol{\lambda}_3''' = \begin{pmatrix} \mathbf{s}_3'''^\top \boldsymbol{\lambda}_3''' \\ v_3 \mathbf{s}_3'''^\top \boldsymbol{\lambda}_3''' \\ w_3 \mathbf{s}_3'''^\top \boldsymbol{\lambda}_3''' \end{pmatrix} \sim \begin{pmatrix} 1 \\ v_3 \\ w_3 \end{pmatrix}$$

Due to lemma 7.1 all three right sides should return the same homogenous vector (i.e. the epipole $\mathbf{e}_1''$). This can only be achieved if the following holds for the coefficients of our parameterization: $v_1 = v_2 = v_3 = v$ and $w_1 = w_2 = w_3 = w$. And this shows that the second four constraints are not satisfied in general and therefore all eight constraints are indeed independent. Consequently by adapting our parameterization in (7.56) we get a new parameterization, that satisfies all eight constraints in general:

$$\mathsf{l}_i = \begin{bmatrix} \mathbf{s}_i''', & v\mathbf{s}_i''' + m_i \mathbf{e}_1''', & w\mathbf{s}_i''' + n_i \mathbf{e}_1''' \end{bmatrix}. \tag{7.57}$$

We will come back to this parameterization later in section 8.3.4. Now we summarize this new minimal set of constraints in the following proposition.

Proposition 7.16 (A new minimal set of constraints via the correlation slices.) *A valence-(1,2) tensor $T_{\cdot\cdot i}^{kj}$ is a trifocal tensor if its correlation slices – the matrices that are obtained by fixing the covariant index i – satisfy the following eight constraints: the three rank constraints* (7.52)*, the epipolar constraint* (7.53) *and the four mapping constraints* (7.55)*.*

The rank constraints are of degree 3 in the elements of the tensor, the epipolar constraint is of degree 6 and the mapping constraints are of degree 7. The latter, however, can be regarded as of degree 5 in practise, as the epipolar lines $\boldsymbol{\rho}_{1i}'''$ in (7.55) are needed only to build lines which are securely different from these epipolar lines and can be regarded as fixed.

A new minimal set derived from the homography slices

Now we consider the homography slices J_j. From table 7.9 on page 100 we see that the ith column of J_j is a point in image Ψ_3, which is the mapping of the intersection of the principal ray $\boldsymbol{\mathcal{R}}_i'$ with the principal plane $\boldsymbol{\Pi}_j''$. Therefore the three columns of each J_j matrix form a triangle J_j in image Ψ_3. From table 7.10 on page 101 and also figure 7.10 on the following page we see that the columns (i.e. the points in image Ψ_3) of the J_j matrices are the same as those of the l_i matrices except for the ordering. Therefore the constraints that we can derive from the columns of the J_j matrices are the same as we found for the columns of the l_i matrices; i.e. the three rank constraints (7.52) and the single epipolar constraint (7.53), which are independent.

As we can see, these four independent constraints is all we can get out of the relations between the points in image Ψ_3. Note that these point relations are valid in $\mathbb{P}^2$ and are therefore *independent* on the actual scaling of the columns of the l_i and J_j matrices. Consequently to get relations that constrain this scale ambiguity we have to consider also the *rows* of the J_j matrices.

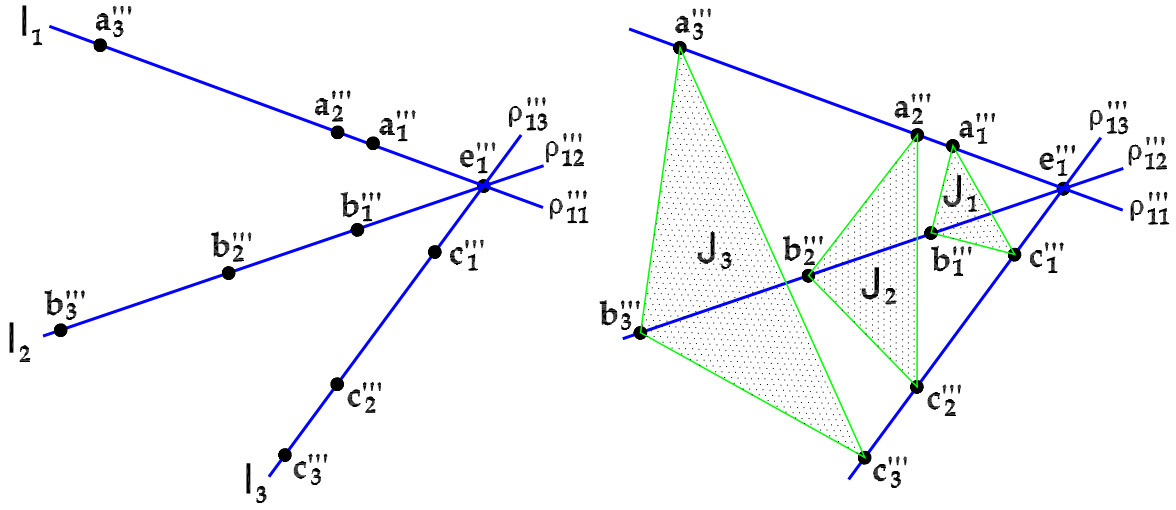

Figure 7.10: *The column vectors of the correlation slices* l_1, l_2 *and* l_3 *(left) and the column vectors of the homography slices* J_1, J_2 *and* J_3 *(right) interpreted as points in image* Ψ_3.

From table 7.9 on page 100 we see that the kth row of J_j is a line in image Ψ_1, which is the mapping of the intersection of the principal plane $\mathbf{\Pi}'''_k$ with the principal plane $\mathbf{\Pi}''_j$. Therefore, the three rows of each J_j matrix form a triangle J_j in image Ψ_1; see figure 7.11.

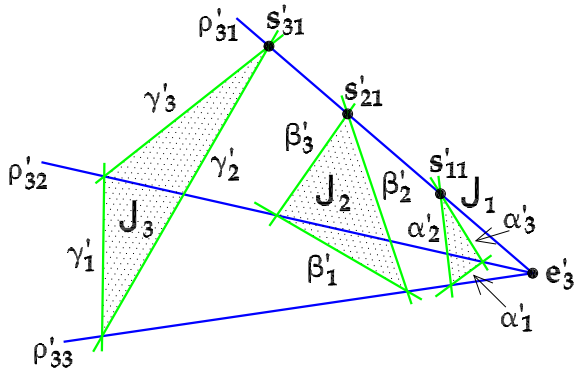

Figure 7.11: *The row vectors of the homography slices* J_1, J_2 *and* J_3 *interpreted as lines in image* Ψ_1*, and three points of intersection* $\mathbf{s}_{j1}$ *which are collinear; cf. figure 7.12.*

Now we consider the corners $\mathbf{s}'_{jk}$ of such a triangle; $\mathbf{s}'_{11}$, $\mathbf{s}'_{21}$ and $\mathbf{s}'_{31}$ are shown in figure 7.11. The corner $\mathbf{s}'_{jk}$ of triangle J_j results from the intersection of the $(k-1)$st and the $(k+1)$st row of J_j. This corner $\mathbf{s}'_{jk}$ is the map of the intersection point $\mathbf{S}_{jk}$ of the three principal planes $\mathbf{\Pi}'''_{k-1}$ and $\mathbf{\Pi}'''_{k+1}$ (both intersecting in the principal ray $\boldsymbol{\mathcal{R}}'''_k$) and $\mathbf{\Pi}''_j$; see figure 7.12 on the next page for the creation of the points $\mathbf{S}_{11}$, $\mathbf{S}_{21}$ and $\mathbf{S}_{31}$. Since $\mathbf{S}_{jk}$ lies on the principal ray $\boldsymbol{\mathcal{R}}'''_k$, the mapped corner $\mathbf{s}'_{jk}$ has to lie on the map of that principal ray, which is the epipolar line $\boldsymbol{\rho}'_{3k}$. This principal ray also intersects the other two principal planes $\mathbf{\Pi}''_{j-1}$ and $\mathbf{\Pi}''_{j+1}$, and so the corners $\mathbf{s}'_{j-1,k}$ and $\mathbf{s}'_{j+1,k}$ arising from the intersection of the $(k-1)$st and $(k+1)$st row of J_{j-1} and J_{j+1} respectively must also lie on $\boldsymbol{\rho}'_{3k}$. So we see that corresponding corners of the three triangles must be collinear – lying on the epipolar lines $\boldsymbol{\rho}'_{3k}$, which are the maps of the principal rays of image

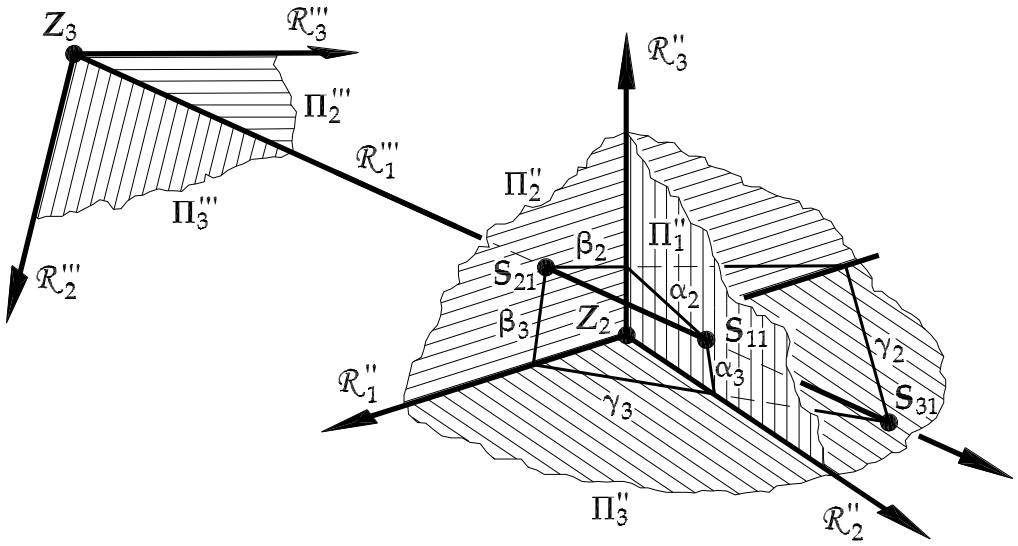

Figure 7.12: *The three points of intersection* $\mathbf{S}_{j1}$ *of the principal ray* $\boldsymbol{\mathcal{R}}'''_1$ *with the principal planes* $\boldsymbol{\Pi}''_j$ *of image* Ψ_2*. The collinearity of the maps of these three points into image* Ψ_1 *gives one constraint.*

Ψ_3. This gives the three following **collinearity constraints**

$$\begin{aligned} &|\boldsymbol{\alpha}'_2 \times \boldsymbol{\alpha}'_3, \quad \boldsymbol{\beta}'_2 \times \boldsymbol{\beta}'_3, \quad \boldsymbol{\gamma}'_2 \times \boldsymbol{\gamma}'_3| = 0 \\ &|\boldsymbol{\alpha}'_3 \times \boldsymbol{\alpha}'_1, \quad \boldsymbol{\beta}'_3 \times \boldsymbol{\beta}'_1, \quad \boldsymbol{\gamma}'_3 \times \boldsymbol{\gamma}'_1| = 0 \\ &|\boldsymbol{\alpha}'_1 \times \boldsymbol{\alpha}'_2, \quad \boldsymbol{\beta}'_1 \times \boldsymbol{\beta}'_2, \quad \boldsymbol{\gamma}'_1 \times \boldsymbol{\gamma}'_2| = 0. \end{aligned} \tag{7.58}$$

Since the principal rays are concurrent in space through $\mathbf{Z}_3$, their maps (the epipolar lines $\boldsymbol{\rho}'_{3k}$) in image Ψ_1 are also concurrent – through the epipole $\mathbf{e}'_3$. And therefore we get another **epipolar constraint**

$$|\boldsymbol{\alpha}'_2 \times \boldsymbol{\alpha}'_3 \times \boldsymbol{\beta}'_2 \times \boldsymbol{\beta}'_3, \quad \boldsymbol{\alpha}'_3 \times \boldsymbol{\alpha}'_1 \times \boldsymbol{\beta}'_3 \times \boldsymbol{\beta}'_1, \quad \boldsymbol{\alpha}'_1 \times \boldsymbol{\alpha}'_2 \times \boldsymbol{\beta}'_1 \times \boldsymbol{\beta}'_2| = 0. \tag{7.59}$$

This constraint was set up with the rows of J_1 and J_2, but in general can be set up with any pair of J_j matrices.

So we easily found another set of four constraints, which are also independent. The latter can be seen again from the fact that the collinearity of the corners of the J_j triangles in general does not enforce the carrier lines (i.e. the epipolar lines $\boldsymbol{\rho}'_{3k}$) to be concurrent. What still needs to be shown is that both groups of constraints that we found for the J_j matrices are independent among themselves.

The independency of both groups can be conjectured from the fact that the first group is derived using the columns and the second group using the rows of J_j matrices. An actual proof by hand seems to be not so easy as it was for the first minimal set derived from the I_i matrices. Therefore the proof was carried out using the software package MATLAB and the respective m-file is attached in appendix E.2.

We summarize this second minimal set of constraints in the following proposition.

Proposition 7.17 (A new minimal set of constraints via the homography slices.) *A valence-(1,2) tensor $T^{kj}_{\cdot\cdot i}$ is a trifocal tensor if its homography slices – the matrices that are obtained by fixing the contra-variant index j – satisfy the following eight constraints: the three rank constraints* (7.52)*, the epipolar constraint* (7.53)*, the three collinearity constraints* (7.58) *and the epipolar constraint* (7.59)*.*

The rank constraints are of degree 3 in the elements of the tensor, the first epipolar constraints is of degree 6, the second one of degree 12 and the collinearity constraints are of degree 6.

Now we are able to prove theorem 7.1 on page 99, which states that the constraints guarantee that the maps of the principal rays in the images are straight and concurrent lines.

Proof:
The three rank constraints (7.52) and the epipolar constraint (7.53) derived from the I_i matrices guarantee that the maps of the principal rays of image Ψ_1 into image Ψ_3 are straight and concurrent lines. These constraints also guarantee that the maps of the principal rays of image Ψ_1 into image Ψ_2 are straight lines, but not necessarily concurrent. However, the latter is guaranteed by the four mapping constraints (7.55) also derived from the I_i matrices. Using these eight constraints, we can parameterize the I_i matrices in the way of (7.57). With this parameterization it is straight forward to show that also the three collinearity constraints (7.58) and the epipolar constraint (7.59) derived from the J_j matrices are satisfied. Therefore, also the maps $\boldsymbol{\rho}'_{3k}$ of the principal rays of image Ψ_3 into image Ψ_1 are straight and concurrent lines.

So, if we look at figure 7.8 on page 99, we find one set of concurrent epipolar lines in each image; i.e. $\boldsymbol{\rho}'_{3k}$, $\boldsymbol{\rho}''_{1i}$ and $\boldsymbol{\rho}'''_{1i}$. What needs to be shown is that the other three sets ($\boldsymbol{\rho}'_{2j}$, $\boldsymbol{\rho}''_{3k}$ and $\boldsymbol{\rho}'''_{2j}$) are also straight and concurrent lines.

Using the theorem of Desargues 3.1 on page 14 this task is rather simple to solve. For example we consider the situation in image Ψ_1. Due to the three collinearity constraints (7.58) and the epipolar constraint (7.59) derived from the J_j matrices we know that the three lines $\boldsymbol{\rho}'_{3k}$ form a set of straight and concurrent lines. The rows of the J_j matrices form triangles J_j and intersect on the lines $\boldsymbol{\rho}'_{3k}$, see figure 7.13 on the next page left part. Since corresponding corners of these triangles lie on the concurrent lines $\boldsymbol{\rho}'_{3k}$, each pair of triangles $\{\mathsf{J}_1, \mathsf{J}_2\}$, $\{\mathsf{J}_1, \mathsf{J}_3\}$ and $\{\mathsf{J}_2, \mathsf{J}_3\}$ is *perspective from the point* $\mathbf{e}'_3$. Due to the theorem of Desargues each pair of triangles $\{\mathsf{J}_{j-1}, \mathsf{J}_{j+1}\}$ is also *perspective from a line* $\boldsymbol{\alpha}'_j$. And since the three pairs of triangles are perspective from the same point $\mathbf{e}'_3$ these three lines $\boldsymbol{\alpha}'_j$ must be concurrent in a point $\mathbf{z}'$.

Now we consider the rows of the K_k matrices. From table 7.9 on page 100 we see that these rows are also lines in image Ψ_1, which therefore make up triangles K_k. And from table 7.10 on page 101 we see that these rows are the same as the rows of the J_j matrices, except for the ordering. Due to this different ordering the sides of the triangle K_k are made up of the kth side of each J_j triangle (see figure 7.13 on the facing page right part).

We know from the J_j triangles that their sides are *perspective from the lines* $\boldsymbol{\alpha}'_j$, which are concurrent in the point $\mathbf{z}'$. Therefore, each pair of triangles $\{\mathsf{K}_{k-1}, \mathsf{K}_{k+1}\}$ must be *perspective from this point* $\mathbf{z}'$. Then, due to the theorem of Desargues, each pair of triangles $\{\mathsf{K}_{k-1}, \mathsf{K}_{k+1}\}$ is also *perspective from a line*, which is the line $\boldsymbol{\rho}'_{3k}$.

The rows of the K_k matrices are related to the maps $\boldsymbol{\rho}'_{2j}$ of the principal rays of image Ψ_2 in image Ψ_1 in the very same way as the J_j matrices are related to the maps $\boldsymbol{\rho}'_{3k}$ of the principal

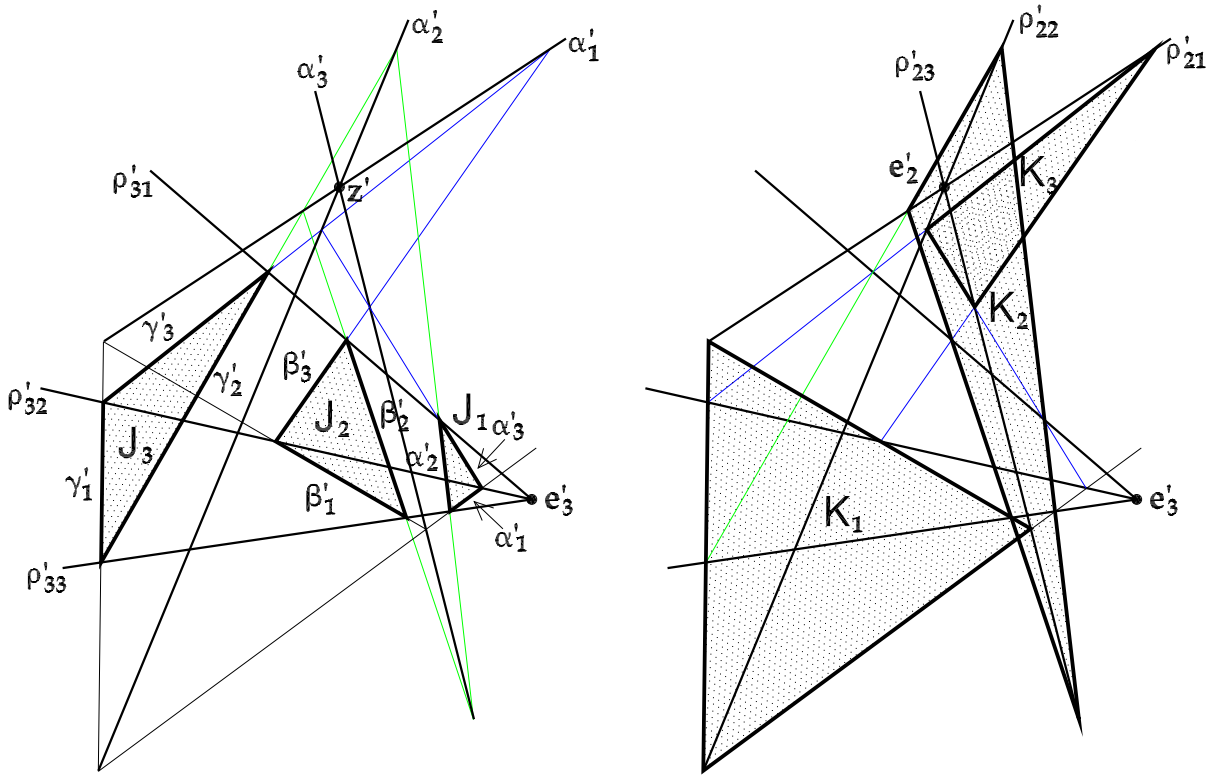

Figure 7.13: *The triangles formed by the rows of the* J_j *matrices (left) and the triangles formed by the rows of the* K_k *matrices (right).*

rays of image Ψ_3 in image Ψ_1; i.e. the sides of the K_k triangles must intersect on the lines $\boldsymbol{\rho}'_{2j}$, and the lines $\boldsymbol{\rho}'_{2j}$ must be concurrent in the epipole $\mathbf{e}'_2$.

From the relations of the K_k triangles to the lines $\boldsymbol{\rho}'_{2j}$ and the relations found using the theorem of Desargues, we see that the lines $\boldsymbol{\alpha}'_j$ are actually identical to the lines $\boldsymbol{\rho}'_{2j}$ and that the point $\mathbf{z}'$ is actually the epipole $\mathbf{e}'_2$.

So we see that given the set of straight and concurrent lines $\boldsymbol{\rho}'_{3k}$ in image Ψ_1, the remaining set of straight and concurrent lines $\boldsymbol{\rho}'_{2j}$ in image Ψ_1 is guaranteed due to the theorem of Desargues. The same reasoning can be applied to image Ψ_2, where we have given the set $\boldsymbol{\rho}''_{1i}$ and require the set $\boldsymbol{\rho}''_{3k}$, and to the image Ψ_3, where we have given the set $\boldsymbol{\rho}'''_{1i}$ and require the set $\boldsymbol{\rho}'''_{2j}$. □

7.7 Historical notes

In photogrammetry Mikhail proposed first to use three images instead of two for the relative orientation of aerial images; [Mikhail 1962]. The benefit gained by using three images is the checking and minimizing of the x-parallaxes, which are uncontrolled in case of two images. This work, however, still relayed on the non-linear collinearity relations (1.1).

In [Rinner and Burkhardt 1972] the dependencies between corresponding points in three images are investigated. These dependencies are termed *trilinear* relations, because corresponding lines in two images are related to a line in the third image. These relations can be used to determine the corresponding point in the third image if the corresponding points in the other two images are given. This way the image content of $n > 3$ images can be transferred into two chosen images to create two images with maximum image content. Consequently the spatial information inherent in all n images could be retrieved from these two distinct images alone. These relations could also be used to form the ground map from two given images – here the ground map acts as the third image.

Despite those and other interesting properties of these trilinear relations they remained quite unused in photogrammetry, maybe due to the lack of a compact mathematical representation. Simple and compact mathematical representations always have been a prime goal in computer vision; and linear representations are the optimum representations in this context.

Historically at the late 1980's only linear methods based on *point correspondences*, introduced by [Longuet-Higgins 1981] in form of the essential matrix for two *calibrated images*, were present in the computer vision community, but no linear methods using lines. Since points and lines are both available from image feature extractors, it was of interest to find also linear methods, that can handle lines.

Motivated by this need, the representation of the relative orientation of three *calibrated images* by means of 27 parameters was discovered independently by [Spetsakis and Aloimonos 1990] and [Weng, Huang, and Ahuja 1992]. Both, works come up with the same two linear constraints for corresponding lines in three images, which correspond to the line-trilinearities (7.23) on page 78. In both work these constraints are expressed using three 3×3 matrices, which correspond to the transposed correlation slices in this thesis. Spetsakis and Aloimonos go even further as they also show that the same three matrices can be used to set up nine constraints for corresponding points in the three images. These nine constraints correspond to the point-trilinearities (7.29) on page 79. Because their work was based on calibrated images, Spetsakis and Aloimonos deduced that only three of these constraints are independent, and therefore 9 point correspondences were required for a direct solution.

In the meantime other authors were also dealing with *uncalibrated images* and tried to solve various vision tasks. Motivated by the problem of predicting a novel view from a set of given ones without reconstructing the 3D scene, Shashua rediscovered the same nine constraints for corresponding points in three uncalibrated images in [Shashua and Wolf 1994] and [Shashua 1995a], which he expressed using nine 3-vectors. Due to the general approach of uncalibrated images these point correspondences were investigated independently and it became evident that due to the expression in these 27 parameters actually 4 of these constraints are independent. Consequently 7 point correspondences are already sufficient to determine these 27 parameters.

Further, Hartley also showed that the methods of [Weng, Huang, and Ahuja 1992] for using lines in calibrated images are equally applicable also for the projective scene reconstruction in the uncalibrated case; [Hartley 1994b]. So far the 27 elements involved in each author's work was realized in different ways grouped either as matrices or vectors. Their first representation as a tensor was given in [Vieville and Luong 1993]. Based on this tensor representation, Hartley showed that the point relations of Shashua and the scene reconstruction from lines arise from the same tensor, for which he proposed the term *trifocal tensor*; [Hartley 1994a].

In subsequent work the tensor and its underlying relations were re-derived using different approaches: [Faugeras and Mourrain 1995] used Grassman-Caley algebra, and [Lasenby and

Lasenby 1998] used Geometric algebra. The properties of the tensor slices were investigated by e.g. [Shashua and Werman 1995] and [Papadopoulo and Faugeras 1998], as were the constraints of the tensor; cf. section 7.6.

In [Avidan and Shashua 1996] the authors consider the relations between all possible trifocal tensors among $n+2$ images, $n > 1$. Using the operators O_{12}, which exchanges the first and the second view, O_{23}, which exchanges the second and the third view, both were already mentioned in section 7.5, and a third transfer operator, which takes the tensors $T(\Psi_i, \Psi_j, \Psi_k)$ and $T(\Psi_j, \Psi_k, \Psi_l)$, and returns the tensor $T(\Psi_i, \Psi_k, \Psi_l)$, all $n(n-1)(n-2)$ tensors among these $n+2$ images can be generated from n arbitrary given tensors of these $n+2$ images.

If only the image triples with identical first and second image are considered, then only n different tensors $T(\Psi_1, \Psi_2, \Psi_i)$, $i = 3, \ldots, n$, exist. Among these n tensors only 12 are linearly independent; [Shashua 1995b]. Consequently any tensor $T(\Psi_1, \Psi_2, \Psi_i)$ can be expressed linearly in 12 basis tensors and due to Shashua the coefficients of this linear combination are determined by 6 corresponding points and the three images Ψ_1, Ψ_2 and Ψ_i.

Chapter 8

The Computation of the Trifocal Tensor

The aim of this chapter is to compute the trifocal tensor for a given set of point and/or line correspondences across three images. The basic equations for this task are the point and line trilinearities presented in section 7.3, which are linear in the 27 elements of the tensor. In principle 7 triples of points or 13 triples of lines are sufficient for a unique solution. Due to the inevitable noise in the data, however, it is advisable to use more correspondences than necessary and to determine the tensor using an adjustment. Since the trilinearities are linear in the 27 elements of the tensor we can in principle determine the trifocal tensor without any difficulties simply by solving a homogenous system of linear equations, just like we did for the fundamental matrix in section 6.1.2 – but now also line correspondences can be used for the computation. This direct linear solution will be discussed in section 8.1. Although this is a fast and easy way to compute the trifocal tensor, it has two mayor drawbacks.

The first drawback is that the trifocal tensor is computed with 26 degrees of freedom, because the 8 intrinsic constraints of section 7.6 are not considered in this direct linear solution. For this reason and the inevitable noise in the data the solution will not be a *valid trifocal tensor*; i.e. one that has 18 degrees of freedom. Furthermore, depending on the number of point and line correspondences, their distribution in space and the arrangement of the three images, chances are that the redundant 8 unconstrained degrees of freedom of the tensor absorb some of the errors in the image measurements and thereby return a severely disturbed tensor and thus a rather wrong relative orientation of the three images.

The second drawback is that this direct linear solution does not minimize the errors in the original measurements (so-called *reprojection error*) but the deviations of the homogenous right sides (so-called *algebraic error*). And therefore we suspect that additional errors in the tensor elements are induced.

From these drawbacks we see that the direct linear solution only provides an initial solution and we should further refine it by enforcing the constraints and by minimizing reprojection error. This can be accomplished by determining the trifocal tensor in the Gauss-Helmert model; see section D. For this we have to consider the parameterization of a valid trifocal tensor in the adjustment. This can be done basically in two ways: By parameterizing the tensor using its 27 elements and introducing 8 suitable constraints from section 7.6 together with the scale fixing; or by using an alternative parameterization for the trifocal tensor that has 18 degrees of freedom. These methods of computing a valid trifocal tensor are discussed in section 8.3.

8.1 The direct linear solution

If we have a set of C image correspondences across three images, which can be any of the following kinds: $\{\boldsymbol{\lambda}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''\}$ $(C_{\lambda\lambda\lambda})$, $\{\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''\}$ $(C_{x\lambda\lambda})$, $\{\mathbf{x}', \mathbf{x}'', \boldsymbol{\lambda}'''\}$ $(C_{xx\lambda})$, $\{\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}'''\}$ $(C_{x\lambda x})$ or $\{\mathbf{x}', \mathbf{x}'', \mathbf{x}'''\}$ (C_{xxx}), where $C_{\ldots}$ gives the number of the respective correspondences so that $C = C_{\lambda\lambda\lambda} + C_{x\lambda\lambda} + C_{xx\lambda} + C_{x\lambda x} + C_{xxx}$. Then we can determine the trifocal tensor in a linear way by solving

$$\boldsymbol{A}\boldsymbol{t} = \mathbf{0}, \tag{8.1}$$

where $\boldsymbol{t}$ is a 27-vector given as

$$\boldsymbol{t} = \begin{pmatrix} \mathrm{vec}\,(\mathsf{I}_1) \\ \mathrm{vec}\,(\mathsf{I}_2) \\ \mathrm{vec}\,(\mathsf{I}_3) \end{pmatrix}. \tag{8.2}$$

The matrix $\boldsymbol{A}$ is the Jacobian of the trilinearities with respect to the tensor elements. It is built row-wise by sub matrices $\boldsymbol{A}_c$ $(c = 1, \ldots, C)$, which are made up of the coefficients of the trilinearities of our image correspondences and which can be expressed using the Kronecker product; see table 8.1.

trilinearities		
entities	$\boldsymbol{A}_c$	n
$\boldsymbol{\lambda}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''$	$\left(\boldsymbol{I}_2 \otimes (\boldsymbol{\lambda}''^{\top} \otimes \boldsymbol{\lambda}'''^{\top})\right)\left(\mathsf{S}^{red}(\boldsymbol{\lambda}') \otimes \boldsymbol{I}_9\right)$	2
$\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''$	$\left(\boldsymbol{\lambda}''^{\top} \otimes \boldsymbol{\lambda}'''^{\top}\right)\left(\mathbf{x}'^{\top} \otimes \boldsymbol{I}_9\right)$	1
$\mathbf{x}', \mathbf{x}'', \boldsymbol{\lambda}'''$	$\left(\mathsf{S}^{red}(\boldsymbol{x}'') \otimes \boldsymbol{\lambda}'''^{\top}\right)\left(\mathbf{x}'^{\top} \otimes \boldsymbol{I}_9\right)$	2
$\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}'''$	$\left(\boldsymbol{\lambda}''^{\top} \otimes \mathsf{S}^{red}(\boldsymbol{x}''')\right)\left(\mathbf{x}'^{\top} \otimes \boldsymbol{I}_9\right)$	2
$\mathbf{x}', \mathbf{x}'', \mathbf{x}'''$	$\left(\mathsf{S}^{red}(\mathbf{x}'') \otimes \mathsf{S}^{red}(\boldsymbol{x}''')\right)\left(\boldsymbol{x}'^{\top} \otimes \boldsymbol{I}_9\right)$	4

Table 8.1: *The sub matrices $\boldsymbol{A}_c$ of coefficients of the reduced trilinearities represented using the Kronecker product; cf. table 7.5 on page 80. The reduced axiator $\mathsf{S}^{red}(\mathbf{x})$ for points is given by relation* (3.48) *on page 28. The last column in this table gives the number n of the linearly independent equations within each group of trilinearities and thus the number of rows in $\boldsymbol{A}_c$.*

Note that although the implementation of the $\boldsymbol{A}_c$ matrices using the Kronecker product is rather simple, it requires a lot of computational overhead. Therefore if computation time is of concern a more direct implementation of the $\boldsymbol{A}_c$ matrices should be aimed for. For example the basic trilinearity for $\{\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''\}$ can simply be written as:

$$\boldsymbol{A}_c(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''') = \left(x'^1\mathbf{h}^{\top},\quad x'^2\mathbf{h}^{\top},\quad x'^3\mathbf{h}^{\top}\right),$$

with

$$\mathbf{h}^{\top} = \left(\lambda''_1\boldsymbol{\lambda}'''^{\top},\quad \lambda''_2\boldsymbol{\lambda}'''^{\top},\quad \lambda''_3\boldsymbol{\lambda}'''^{\top}\right).$$

In general, the system (8.1) has only a unique solution for the homogenous trifocal tensor incorporated in $\boldsymbol{t}$ if exactly 26 equations are used. However, as such minimal solutions are rather unstable, we will use many more equations usually and solve system (8.1) by minimizing algebraic error (i.e. the squared sum of the right side of the equations)

$$\boldsymbol{t}^\top \mathsf{A}^\top \mathsf{A} \boldsymbol{t} \;\rightarrow\; \min \quad \text{with} \quad ||\boldsymbol{t}|| = 1. \tag{8.3}$$

Again as in the case of the fundamental matrix, the solution is found by computing a singular value decomposition of A. An alternative solution without the SVD would be to set the ith element of $\boldsymbol{t}$ to 1 and to put the respective column of A on the right side of (8.1), thereby creating an inhomogeneous system of linear equations:

$$\mathsf{A}_i \boldsymbol{t}_i = -\boldsymbol{a}_i \quad \rightarrow \quad \boldsymbol{t}_i = -(\mathsf{A}_i^\top \mathsf{A}_i)^{-1} \mathsf{A}_i^\top \boldsymbol{a}_i,$$

where A_i and $\boldsymbol{t}_i$ represent the matrix and vector obtained by striking out the ith column and ith element respectively. Note, however, that the solution obtained in this way will depend on the value of i. This is due to the algebraic error, which is different for each choice of i, and the fact that the original matrix A is *regular* for noisy data. Therefore, it is advisable to use the SVD solution when minimizing algebraic error, as in this case all columns of A contribute equally to the solution.

As it was already mentioned in the beginning of this chapter, the direct linear solution for the trifocal tensor has two disadvantages. First, the computed tensor is not valid, because it is computed with 26 degrees of freedom without considering the constraints of section 7.6. And second, this direct linear solution does not minimize reprojection error but algebraic error. Therefore we will consider in the following sections, how to overcome these two disadvantages.

8.2 Minimizing reprojection error

For determining the trifocal tensor by minimizing reprojection error we can use the *Gauss-Helmert model*, which is described in detail in section D.5 on page 195. This model is based on linearized equations, and therefore the estimation of the tensor $\boldsymbol{t}$ is iterative, starting from an initial guess $\boldsymbol{t}^0$. The latter may be obtained from the direct linear solution for example. In the Gauß-Helmert model we have to distinguish between two different types of equations: (i) *conditions*, which are relations involving the observations and the unknown tensor elements, and (ii) *constraints*, which are relations involving only the tensor elements. For the trifocal tensor the conditions are the trilinearities and if the internal constraints are not considered, then there is only one constraint, which is the fixing of the tensor scale e.g. by setting its Frobenius norm to 1. Basically the Gauss-Helmert model can be written in the following way:

$$\begin{aligned} &\text{conditions: } \mathsf{A}\Delta\boldsymbol{t} + \mathsf{B}\boldsymbol{v} - \boldsymbol{w} = \boldsymbol{0}, \\ &\text{constraints: } \mathsf{C}\Delta\boldsymbol{t} - \boldsymbol{d} = \boldsymbol{0}, \end{aligned}$$

where $\Delta\boldsymbol{t}$ is the vector of corrections for the unknown tensor elements, $\boldsymbol{v}$ is the residual vector and contains the reprojection error of the observed image features, A is the Jacobian of the trilinearities with respect to the tensor elements, B is the Jacobian of the trilinearities with respect to the observations, and $\boldsymbol{w}$ is the contradiction of the trilinearities. The matrix C is

the Jacobian of the constraint(s) with respect to the tensor elements, and $\boldsymbol{d}$ is the respective contradiction. In case of *disregarding* the internal constraints, only the scaling constraint, e.g. $|\boldsymbol{t}| = 1$, is involved and $\boldsymbol{C}$ becomes a row vector.

The solution for the trifocal tensor in the Gauß-Helmert model is found by minimizing the weighted sum of the squared residuals:

$$\boldsymbol{v}^\top \boldsymbol{P} \boldsymbol{v} \quad \rightarrow \quad \min.$$

The weight matrix of the observations $\boldsymbol{P}$ is proportional to the inverse of the covariance matrix of the observed image features Σ by an unknown scale factor σ_0^2, the reference variance: $\boldsymbol{P} = \sigma_0^2 \Sigma^{-1}$; see section D.3 on page 191. The corrections $\Delta\boldsymbol{t}$ are found by solving the following system

$$\begin{bmatrix} \boldsymbol{A}^\top \boldsymbol{W} \boldsymbol{A} & \boldsymbol{C}^\top \\ \boldsymbol{C} & \boldsymbol{0} \end{bmatrix} \begin{pmatrix} \Delta\boldsymbol{t} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}^\top \boldsymbol{W} \boldsymbol{w} \\ \boldsymbol{d} \end{pmatrix}, \tag{8.4}$$

with $\boldsymbol{\mu}$ being a vector of Lagrange multipliers (see section D.5). The matrix $\boldsymbol{W}$ is the weight matrix of the trilinearities obtained by error propagation as

$$\boldsymbol{W} = (\boldsymbol{B} \boldsymbol{P}^{-1} \boldsymbol{B}^\top)^{-1}.$$

From the comparison of the algebraic minimization (8.3) with (8.4) the reason for the weakness of the algebraic solution becomes evident: The trilinearities are not correctly weighted during the algebraic minimization, as there implicitly $\boldsymbol{W} = \boldsymbol{I}$ is used.

The Jacobian $\boldsymbol{A}$ is build row-wise by sub Jacobian matrices $\boldsymbol{A}_c$ $(c = 1, \ldots, C)$ for each type of trilinearity. These sub matrices are identical to the ones presented in table 8.1 on page 112. Analogously, the Jacobian $\boldsymbol{B}$ is build row-wise by sub Jacobian matrices $\boldsymbol{B}_c$ for each type of trilinearity. In table 8.2 on the facing page the sub Jacobian matrices for the two most important trilinearities $T(\mathbf{x}', \mathbf{x}'', \mathbf{x}''') = 0$ and $T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''') = 0$ are given. Note, that for the point trilinearities $T(\mathbf{x}', \mathbf{x}'', \mathbf{x}''') = 0$ the rank of the respective sub Jacobian depends on the degrees of freedom of the trifocal tensor. If the tensor is computed unconstrained with > 18 degrees of freedom the rank is 4, whereas for the computation of a valid trifocal tensor with 18 degrees of freedom the rank is 3; [Förstner 2000]. The latter is easy to see from the fact that a triple of corresponding image points is made up of 6 coordinates, the corresponding point in object space has 3 unknown coordinates, which gives a redundancy of 3.

This has two consequences: First, the minimum number of point correspondences across three images required for computing a *valid* trifocal tensor is 6 not 7. For this computation an initial tensor is needed, which can be obtained using a 7th point and the direct linear solution. In section 8.3.2 a direct method for computing a valid trifocal tensor from 6 point correspondences is presented, which, however, may return up to three solutions. The second consequence when computing a valid trifocal tensor using point correspondences is, that we may only consider 3 of the 4 reduced point trilinearities. In [Förstner 2000] ways are presented to choose 3 relations in an optimal way. The other possibility is to use the reduced 4 trilinearities as in the unconstrained case and to enforce the correct rank by applying an SVD for each sub Jacobian matrix $\boldsymbol{B}_c$.

During several tests of computing an *unconstrained* trifocal tensor by minimizing reprojection error in the Gauß-Helmert model using synthetic point data with added normally distributed noise the following behavior was observed. The iterative solution was initialized by the direct

$$T(\mathbf{x}', \mathbf{x}'', \mathbf{x}''') = \mathbf{0}$$

$$\underset{4\times 6}{\mathsf{B}_c} = \left[\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial x'}, \frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial y'}, \frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial \boldsymbol{x}''}, \frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial \boldsymbol{x}'''}\right] \text{ with } \operatorname{rank}(\mathsf{B}_c) = 3 \text{ or } 4$$

$$\begin{aligned}
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial x'} &= \operatorname{vec}\left(\mathsf{S}^{red}(\mathbf{x}''')\mathsf{l}_1\mathsf{S}^{red\top}(\mathbf{x}'')\right) \\
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial y'} &= \operatorname{vec}\left(\mathsf{S}^{red}(\mathbf{x}''')\mathsf{l}_2\mathsf{S}^{red\top}(\mathbf{x}'')\right) \\
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial \boldsymbol{x}''} &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \left(\mathsf{S}^{red}(\mathbf{x}''')\mathsf{J}_3\mathbf{x}'\right) \\
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''')}{\partial \boldsymbol{x}'''} &= \left(\mathsf{S}^{red}(\mathbf{x}'')\mathsf{K}_3\mathbf{x}'\right) \otimes \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}
\end{aligned}$$

$$T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''') = 0$$

$$\underset{1\times 10}{\mathsf{B}_c} = \left[\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}'}, \frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}''_S}, \frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}''_E}, \frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}'''_S}, \frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}'''_E}\right] \text{ with } \operatorname{rank}(\mathsf{B}_c) = 1$$

$$\begin{aligned}
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}'} &= \mathbf{u}_1^\top \bar{I} \\
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}''_S} &= \mathbf{u}_2^\top \mathsf{S}(\mathbf{x}''_E)\bar{I} \\
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}''_E} &= -\mathbf{u}_2^\top \mathsf{S}(\mathbf{x}''_S)\bar{I} \\
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}'''_S} &= \mathbf{u}_3^\top \mathsf{S}(\mathbf{x}'''_E)\bar{I} \\
\frac{\partial T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''')}{\partial \boldsymbol{x}'''_E} &= -\mathbf{u}_3^\top \mathsf{S}(\mathbf{x}'''_S)\bar{I}
\end{aligned}$$

with:

$$\begin{aligned}
\boldsymbol{\lambda}'' &= \frac{\mathbf{x}''_S \times \mathbf{x}''_E}{||\mathbf{x}''_S \times \mathbf{x}''_E||} \\
\boldsymbol{\lambda}''' &= \frac{\mathbf{x}'''_S \times \mathbf{x}'''_E}{||\mathbf{x}'''_S \times \mathbf{x}'''_E||} \\
\bar{I} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\
\mathbf{u}_1^\top &= \boldsymbol{\lambda}'''^\top \begin{bmatrix} \mathsf{l}_1\boldsymbol{\lambda}'', & \mathsf{l}_2\boldsymbol{\lambda}'', & \mathsf{l}_3\boldsymbol{\lambda}'' \end{bmatrix} \\
\mathbf{u}_2^\top &= \boldsymbol{\lambda}'''^\top \begin{bmatrix} \mathsf{J}_1\mathbf{x}', & \mathsf{J}_2\mathbf{x}', & \mathsf{J}_3\mathbf{x}' \end{bmatrix} \frac{1}{||\mathbf{x}''_S \times \mathbf{x}''_E||^3} \mathsf{S}(\mathbf{x}''_S \times \mathbf{x}''_E)\mathsf{S}(\mathbf{x}''_S \times \mathbf{x}''_E) \\
\mathbf{u}_3^\top &= \boldsymbol{\lambda}''^\top \begin{bmatrix} \mathsf{K}_1\mathbf{x}', & \mathsf{K}_2\mathbf{x}', & \mathsf{K}_3\mathbf{x}' \end{bmatrix} \frac{1}{||\mathbf{x}'''_S \times \mathbf{x}'''_E||^3} \mathsf{S}(\mathbf{x}'''_S \times \mathbf{x}'''_E)\mathsf{S}(\mathbf{x}'''_S \times \mathbf{x}'''_E)
\end{aligned}$$

Table 8.2: *The sub Jacobian matrices B_c of the trilinearities $T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''') = \mathbf{0}$ and $T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''') = 0$ with respect to the observations $\mathbf{x}' = (x', y', 1)^\top = (\boldsymbol{x}'^\top, 1)^\top$, $\mathbf{x}'' = (\boldsymbol{x}''^\top, 1)^\top$, $\mathbf{x}''' = (\boldsymbol{x}'''^\top, 1)^\top$, $\mathbf{x}''_{S,E} = (\boldsymbol{x}''^\top_{S,E}, 1)^\top$ and $\mathbf{x}'''_{S,E} = (\boldsymbol{x}'''^\top_{S,E}, 1)^\top$. For the trilinearity $T(\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}''') = 0$ it is assumed that each line is defined by its start and end point. The reduced axiator $\mathsf{S}^{red}(\mathbf{x})$ for points is given by relation* (3.48) *on page 28. The rank of B_c of the point trilinearity $T(\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}''') = \mathbf{0}$ is 4 if the trifocal tensor is computed unconstrained with > 18 degrees of freedom, and the rank is 3 if the tensor is computed with 18 degrees of freedom.*

linear solution and therefore the rank of the sub Jacobian matrices B_c of the point trilinearities was 4. At the very end of the iterations the rank should still be 4 as the internal constraints were not considered. At the end, however, the rank turned out to be numerically near to 3; i.e. the respective unknowns correspond almost to a valid tensor. Because of this rank defect of almost 1 in B_c the inversion of W in (8.4) induced wrong large corrections in the unknowns, which destroyed the quality of the previously determined tensor. Consequently the rank of B_c became 4 again and the iteration proceeded again till this break point was reached another time. This way an oscillating state was reached.

The explanation for this behavior seems to lie in the synthetic and therefore perfectly normally distributed noise in the test data. A certain part of this noise was eliminated due to using the Gauß-Helmert model and minimizing reprojection error. Consequently the solution of the tensor is near to a valid tensor, with a distance depending on the remaining noise in the data. Therefore the rank of the sub Jacobians B_c tends to become 3 the more the iteration approaches its final state.

8.3 The constrained solution

In this section we consider the computation of a valid trifocal tensor; i.e. one that satisfies the internal constraints and therefore has 18 degrees of freedom. We can distinguish the methods for computing a valid trifocal tensor in two aspects.

The first aspect concerns the *realization of the 18 degrees of freedom*. This can be done either by

(A) introducing a set of sufficient constraints into the computation, or by

(B) introducing an alternative parameterization for the trifocal tensor, which has exactly 18 degrees of freedom.

The second aspect concerns the *input for the constrained solution*. This can either be

(a) the original measurements, or

(b) the direct linear solution of the trifocal tensor.

In the case (a) we compute that constrained solution, which fits best to the original point and line measurements in the images, by minimizing algebraic or reprojection error. And in case (b) we compute that constrained solution, which lies closest to the direct linear solution, by minimizing some type of algebraic error. In [Faugeras and Luong 2001] this method is termed *projection method*, as the given direct linear solution is projected orthogonally onto the manifold defined by the constraints of the trifocal tensor. Since only the case (a) allows minimizing of reprojection error, this case is the preferred one.

Note that all methods for computing a valid trifocal tensor are nonlinear, regardless of the realization of the 18 degrees of freedom and the input for the constrained solution. This follows from the fact that all internal constraints presented in section 7.6 are nonlinear in the elements of the tensor. Consequently, this nonlinear solution requires an initialization, which in general is provided by the direct linear solution.

In the following we consider the first aspect and show ways of introducing the 18 degrees of freedom in our computation. For case (A) we can use any of the sufficient sets of constraints presented in section 7.6. For case (B) several parameterizations having 18 degrees were proposed in the past and we will review the most important ones in the following.

8.3.1 The parameterization using the projection matrices

Since the trifocal tensor describes the relative orientation of three images, it is a natural approach to parameterize the tensor in the elements of the three projection matrices. And as we saw in section 7.4.3 on page 90, the first projection matrix can be set to the standard form

$$\mathsf{P}_1 = [I_3, \mathbf{0}]\,,$$

as this just corresponds to a projective transformation in the object space, which does not affect the relative orientation of the images. We further saw that the other two projection matrices can then be represented as

$$\begin{aligned}\mathsf{P}_2 &\sim [\mathsf{A}, \mathbf{e}_1'']\,,\\ \mathsf{P}_3 &\sim [\mathsf{B}, \mathbf{e}_1''']\,.\end{aligned}$$

And using the tensor relation (7.7) on page 63 we therefore can parameterize the tensor in the elements of P_2 and P_3 as

$$T_{\cdot\cdot i}^{kj} \sim e_1''^{\,j} B_{\cdot i}^{k} - e_1'''^{\,k} A_{\cdot i}^{j}. \tag{8.5}$$

Although this parameterization uses 24 parameters, 6 more than required, it is *consistent*; i.e. any choice of the 24 parameters returns a valid trifocal tensor. It is a bilinear parameterization, which can be initialized by the direct linear solution and the method for retrieving the projection matrices described in section 7.4.3. Recall that this method of retrieving valid projection matrices relies on the fact that neither the second nor the third projection center coincides with the first one; therefore in case two of the three projection centers coincide, the roles of the images have to be changed so that the image with the unique projection center plays the role of the first image.

This parameterization was proposed by Hartley in [Hartley 1994a]. And in [Hartley and Zisserman 2001] this parameterization is used to compute the trifocal tensor in different ways: by minimizing algebraic error with fixed or varying epipoles, and by minimizing reprojection error. For these least squares estimations either the SVD or the Levenberg-Marquard algorithm is used. These two methods do not require the introduction of additional constraints to cope with the rank defect due to the redundant 6 parameters, whereas other least squares solvers would need them, especially the Gauß-Helmert model.

The 6 constraints required for the Gauß-Helmert model are explained in the following. Since projection matrices are defined only up to scale, two constraints are required to fix the scale of P_2 and P_3; e.g. by setting $|\mathbf{e}_1''| = 1$ and $|\mathbf{e}_1'''| = 1$. In section 7.4.3 on page 90 we saw that the projective transformation, which takes P_1 to the standard form $[I_3, \mathbf{0}]$ is only defined up to four degrees of freedom – the vector $\mathbf{v}$ and the scalar w in section 7.4.3.

The remaining constraints fix these degrees of freedom. Following [Hartley 1994a] the vector $\mathbf{v}$ is fixed by requiring that the three columns of A are orthogonal to $\mathbf{e}_1''$. This relation between

A and $\mathbf{e}_1''$ is also the basis for retrieving the projection matrices in section 7.4.3. The last constraint fixes w, which acts as a scale of the trifocal tensor. Therefore w can be fixed by setting $||T^{kj}_{\cdot\cdot i}||_{\text{Frob}} = 1$ or in a simpler form by setting $||\mathsf{A}||_{\text{Frob}} = 1$. To understand the latter, consider the tensor $T^{kj}_{\cdot\cdot i}$ with $||T^{kj}_{\cdot\cdot i}||_{\text{Frob}} = 1$. In this case $||\mathsf{A}||_{\text{Frob}} = a \neq 1$. Since the tensor is homogenous, we can divide it by a. Consequently $||T^{kj}_{\cdot\cdot i}||_{\text{Frob}}$ changes to $1/a$ and $||\mathsf{A}||_{\text{Frob}}$ to 1.

If this parameterization is to be used in the Gauß-Helmert model introduced in section 8.2, then the unknowns are not the 27 tensor elements but the 24 elements of the two projection matrices P_2 and P_3. Consequently the Jacobian matrix $\boldsymbol{A}$ of the Gauß-Helmert model, which holds the derivatives of the trilinearities with respect to the unknowns, must be adapted.

The 27 tensor elements make up the vector $\boldsymbol{t}$ introduced in equation (8.2) on page 112

$$\boldsymbol{t} = \begin{pmatrix} \operatorname{vec}(\mathsf{I}_1) \\ \operatorname{vec}(\mathsf{I}_2) \\ \operatorname{vec}(\mathsf{I}_3) \end{pmatrix},$$

where I_i are the correlation slices; see relation 7.10 on page 66. We can put the 24 elements of the parameterization using the projection matrices into a vector $\boldsymbol{p}$ as

$$\boldsymbol{p} = \begin{pmatrix} \operatorname{vec}(\boldsymbol{A}) \\ \operatorname{vec}(\boldsymbol{B}) \\ \mathbf{e}_1'' \\ \mathbf{e}_1''' \end{pmatrix}. \tag{8.6}$$

With the vectors $\boldsymbol{t}$ and $\boldsymbol{p}$ we can represent the derivative of any trilinearity $T(\cdot,\cdot,\cdot)$ as

$$\frac{\partial T(\cdot,\cdot,\cdot)}{\partial \boldsymbol{p}} = \frac{\partial T(\cdot,\cdot,\cdot)}{\partial \boldsymbol{t}} \frac{\partial \boldsymbol{t}}{\partial \boldsymbol{p}}.$$

If we further represent the respective Jacobian matrices by $\boldsymbol{A}_t$ and $\boldsymbol{A}_p$, where $\boldsymbol{A}_t$ follows from the trilinear derivatives of table 8.1 on page 112, then $\boldsymbol{A}_p$ is obtained as

$$\boldsymbol{A}_p = \boldsymbol{A}_t \boldsymbol{D}_p, \tag{8.7}$$

with

$$\underset{27\times 24}{\boldsymbol{D}_p} = \frac{\partial \boldsymbol{t}}{\partial \boldsymbol{p}} = \left[\begin{array}{c|c|c|c} -\boldsymbol{I}\otimes\mathbf{e}_1''' & \mathbf{e}_1''\otimes\boldsymbol{I} & \boldsymbol{I}\otimes\mathbf{b}_1 & -\mathbf{a}_1\otimes\boldsymbol{I} \\ \hline -\boldsymbol{I}\otimes\mathbf{e}_1''' & \mathbf{e}_1''\otimes\boldsymbol{I} & \boldsymbol{I}\otimes\mathbf{b}_2 & -\mathbf{a}_2\otimes\boldsymbol{I} \\ \hline -\boldsymbol{I}\otimes\mathbf{e}_1''' & \mathbf{e}_1''\otimes\boldsymbol{I} & \boldsymbol{I}\otimes\mathbf{b}_3 & -\mathbf{a}_3\otimes\boldsymbol{I} \end{array}\right], \tag{8.8}$$

where the positions of zero elements have been left blank for clearer representation. The identity matrix of dimension 3 is represented as $\boldsymbol{I}$ and the column vectors of A and B are given as $\mathbf{a}_i$ and $\mathbf{b}_i$ respectively.

8.3.2 The six-point parameterization

Torr and Zisserman presented an interesting method to compute a valid trifocal tensor given the minimum number of six point correspondences across three images, which may return up

to three real solutions; [Torr and Zisserman 1997]. This method, which is based on earlier works of Quan [Quan 1994] and Weinshall et al. [Weinshall, Werman, and Shashua 1996], is outlined in the following.

Let the three projection matrices be denoted by P_i and the six space points by $\mathbf{X}_j$. Since the absolute orientation is of no concern we can choose the first five of the space points to act as a projective canonical basis: $\mathbf{X}_1 \sim (1,0,0,0)^\top$, $\mathbf{X}_2 \sim (0,1,0,0)^\top$, $\mathbf{X}_3 \sim (0,0,1,0)^\top$, $\mathbf{X}_4 \sim (0,0,0,1)^\top$ and $\mathbf{X}_5 \sim (1,1,1,1)^\top$. The sixth point in this canonical basis is then given by $\mathbf{X}_6 \sim (P,Q,R,S)^\top$. The corresponding image points $\mathbf{x}_j^i$ (for simpler notation we drop the bars ($'$) and use the superscript i instead) in the three images can be transformed by a suitable projective collineation H_i in $\mathbb{P}^2$ from their original coordinates to the canonical points in $\mathbb{P}^2$: $\mathbf{x}_1^i \sim (1,0,0)^\top$, $\mathbf{x}_2^i \sim (0,1,0)^\top$, $\mathbf{x}_3^i \sim (0,0,1)^\top$ and $\mathbf{x}_4^i \sim (1,1,1)^\top$. In these canonical bases the maps of the fifth and sixth object point are given by $\mathbf{x}_5^i \sim (u_5^i, v_5^i, w_5^i)^\top$ and $\mathbf{x}_6^i \sim (u_6^i, v_6^i, w_6^i)^\top$.

With this preparation it is straight forward to see from $\mathbf{x}_j^i \sim \mathsf{P}_i \mathbf{X}_j$, $j = \{1,2,3,4\}$, that the projection matrices P_i can be expressed as

$$\mathsf{P}_i \sim \begin{bmatrix} \alpha^i & 0 & 0 & \delta^i \\ 0 & \beta^i & 0 & \delta^i \\ 0 & 0 & \gamma^i & \delta^i \end{bmatrix}.$$

If the three projection matrices are known, then also the trifocal tensor is determined. Therefore the main task is to determine the four elements of each projection matrix, which is done by means of the fifth and sixth point. The rest of the algorithm's outline is due to [Faugeras and Luong 2001].

Using the maps $\mathbf{x}_5^i \sim (u_5^i, v_5^i, w_5^i)^\top$ of the fifth point $\mathbf{X}_5 \sim (1,1,1,1)^\top$ we can express the elements of the projection matrices as

$$\begin{aligned} \alpha^i &= u_5^i - \delta^i, \\ \beta^i &= v_5^i - \delta^i, \\ \gamma^i &= w_5^i - \delta^i. \end{aligned}$$

This also fixes the scale of the projection matrices, observe the '=' sign. Since the maps of the sixth point $\mathbf{X}_6 \sim (P,Q,R,S)^\top$ are known to be $\mathbf{x}_6^i \sim (u_6^i, v_6^i, w_6^i)^\top$ we can set up the identity check in $\mathbb{P}^2$:

$$\mathbf{x}_6^i \times (\mathsf{P}_i \mathbf{X}_6) = \mathbf{0},$$

which gives

$$\mathbf{x}_6^i \times \left(\operatorname{diag}(P,Q,R) \begin{pmatrix} u_5^i - \delta^i \\ v_5^i - \delta^i \\ w_5^i - \delta^i \end{pmatrix} + \delta^i \begin{pmatrix} S \\ S \\ S \end{pmatrix} \right) = \mathbf{x}_6^i \times \begin{pmatrix} P\,u_5^i \\ Q\,v_5^i \\ R\,w_5^i \end{pmatrix} + \delta^i \mathbf{x}_6^i \times \begin{pmatrix} S-P \\ S-Q \\ S-R \end{pmatrix} = \mathbf{0}.$$

If $\mathbf{X}_6 \sim (P,Q,R,S)^\top$ were known, we could determine δ^i from this relation. For this the other two vectors must be parallel and therefore it must hold

$$\left(\mathbf{x}_6^i \times \begin{pmatrix} P\,u_5^i \\ Q\,v_5^i \\ R\,w_5^i \end{pmatrix} \right) \times \left(\mathbf{x}_6^i \times \begin{pmatrix} S-P \\ S-Q \\ S-R \end{pmatrix} \right) = \mathbf{0}.$$

Expanding this cross product using relation (B.18) on page 179 we get

$$\begin{vmatrix} u_6^i & P\,u_5^i & S-P \\ v_6^i & Q\,v_5^i & S-Q \\ w_6^i & R\,w_5^i & S-R \end{vmatrix} \mathbf{x}_6^i = 0.$$

Since $\mathbf{x}_6^i \neq \mathbf{0}$ the determinant in the previous equation must be zero. By expanding this determinant we get

$$\begin{aligned} &u_5^i(v_6^i - w_6^i)PS + v_5^i(w_6^i - u_6^i)QS + w_5^i(u_6^i - v_6^i)RS + \\ &u_6^i(v_5^i - w_5^i)QR + v_6^i(w_5^i - u_5^i)PR + w_6^i(u_5^i - v_5^i)PQ = 0. \end{aligned}$$

This equation can be set up for each image and therefore we get a homogenous system of 3 linear equations for the 6-dimensional vector $\mathbf{U} = (PS, QS, RS, QR, PR, PQ)^\top$ of the form $M\mathbf{U} = \mathbf{0}$. Consequently the general solution of this system (or the nullspace of the respective matrix) is 3-dimensional, and since the sum of the coefficients in the previous equation is zero this solution can be expressed as

$$\mathbf{U} \sim \mathbf{A} + \lambda_1\mathbf{B} + \lambda_2(1,1,1,1,1,1)^\top,$$

where $\mathbf{A}$ and $\mathbf{B}$ can be computed using the singular value decomposition of M.

The homogenous vector $\mathbf{U}$ depends only on the coordinates of the sixth object point $\mathbf{X}_6$, which has 3 degrees of freedom. Therefore the elements of $\mathbf{U}$ have to satisfy 2 additional constraints: $U_1U_4 = U_2U_5 = U_3U_6$. These two constraints together with the general representation of the solution $\mathbf{U}$ gives the following two linear equations in λ_2

$$\begin{aligned} &((B_1 + B_4 - B_3 - B_6)\lambda_1 + A_1 + A_4 - A_3 - A_6)\lambda_2 + \\ &((B_1B_4 - B_3B_6)\lambda_1^2 + (A_1B_4 - A_3B_6 + B_1A_4 - B_3A_6)\lambda_1 + A_1A_4 - A_3A_6) = 0, \\ &((B_1 + B_4 - B_2 - B_5)\lambda_1 + A_1 + A_4 - A_2 - A_5)\lambda_2 + \\ &((B_1B_4 - B_2B_5)\lambda_1^2 + (A_1B_4 - A_2B_5 + B_1A_4 - B_2A_5)\lambda_1 + A_1A_4 - A_2A_5) = 0. \end{aligned}$$

By eliminating λ_2 the following equation of degree 3 in λ_1 is obtained

$$\begin{aligned} &((B_1 + B_4 - B_2 - B_5)\lambda_1 + A_1 + A_4 - A_2 - A_5) \\ &((B_1B_4 - B_3B_6)\lambda_1^2 + (A_1B_4 - A_3B_6 + B_1A_4 - B_3A_6)\lambda_1 + A_1A_4 - A_3A_6) - \\ &((B_1 + B_4 - B_3 - B_6)\lambda_1 + A_1 + A_4 - A_3 - A_6) \\ &((B_1B_4 - B_2B_5)\lambda_1^2 + (A_1B_4 - A_2B_5 + B_1A_4 - B_2A_5)\lambda_1 + A_1A_4 - A_2A_5) = 0. \end{aligned}$$

This cubic equation may yield up to three real solutions for λ_1 and each of them provides unique solutions for λ_2, $\mathbf{U}$ (and therefore $\mathbf{X}_6$), δ^i and $\alpha^i, \beta^i, \gamma^i$. So we may get up to three different sets of projection matrices P_i and consequently up to three different trifocal tensors using the relations of section 7.1.1 on page 65.

This 6-point approach is useful in two ways. First, it can be used as a search engine for the so-called *random sample consensus paradigm* (RANSAC), [Fischler and Bolles 1981], which is a useful algorithm for finding gross errors in the data. And second, it can be used to form a *minimal parameterization* for the trifocal tensor. A minimal parameterization has the same number of parameters as the degrees of freedom of the tensor; i.e. 18.

RANSAC. As opposed to common adjustment techniques, which use all data to find an initial solution and remove outliers afterwards, the idea behind RANSAC is to select a random sample of minimum data required to obtain a solution (e.g. two points for a 2D-line) and to test the other data if it fits with the obtained minimal solution. Depending on some threshold the other data can therefore be classified as in- and outliers with respect to the selected sample. This way a consensus measure (the number of outliers) is assigned to the selected sample. This process is repeated enough times to assure with some probability α that at least one randomly chosen sample contains only inliers. The required number n of samples depends on the probability α, the possible percentage of outliers p and the sample size s, and is given by $n \geq \ln(1-\alpha)/\ln(1-(1-p)^s)$. In the meantime several adaptions of the basic RANSAC idea have been published; see e.g. [Tordoff and Murray 2002].

To use RANSAC for the computation of the trifocal tensor using data that is probably corrupted by gross errors, we only have to provide tow things: (i) An algorithm, that works for the minimum point number; like the presented 6-point algorithm. And (ii) a test measure with proper threshold, which can be used to decide if a point triple is an in- or outlier; e.g. the mean error of the spatial intersection for each point triple using the projection matrices obtained from the 6 point sample. For more details see the original paper [Torr and Zisserman 1997].

Note that in principle also the direct linear solution of section 8.1, which requires 7 point triples, can be used for RANSAC. With this algorithm, however, the sample size is larger by one and thus following from the formula above, more samples are needed to assure the same probability for an outlier-free sample, and further the direct linear solution does not return a valid trifocal tensor.

Minimal parameterization. This 6-point approach also provides a possible minimal parameterization for the trifocal tensor, which has 18 parameters and consequently any realization of these parameters returns a valid trifocal tensor. This minimal parameterization is obtained in the following simple way. The 6 triples of corresponding image points have in total 36 coordinates. Therefore if we fix 18 convenient coordinates of them and allow the other 18 to be free, we get the trifocal tensor functionally dependent on the free 18 coordinates; i.e. 18 degrees of freedom.

In this way we can compute a valid trifocal tensor from *all* point correspondences (after removing the outliers). However, it is not obvious from the beginning which 18 coordinates to keep fix and which to consider free. In the original paper [Torr and Zisserman 1997] different choices of 18 coordinates are considered, as well as choices with up to 30 free coordinates. In the latter case the parameterization is no longer minimal but still consistent, and due to this overparameterization a special minimization algorithm by [Gill and Murray 1978] needs to be applied, that 'discards redundant combinations of the variables'. The experiments of Torr and Zisserman show that the final results do indeed depend on the selected coordinates and also the convergence (indicated in the number of iterations) varies largely. The computational effort seems to be further increased by the fact that the usage of the 6-point algorithm is linked to the particular choice of canonical base points in the images. Therefore after each iteration when the parameters (i.e. the free coordinates) have changed, these base points must be updated.

8.3.3 The minimal parameterization by Papadopoulo and Faugeras

The minimal parameterization proposed by Papadopoulo and Faugeras is based on the observation that the correlation slices l_i are always singular matrices with $\operatorname{rank}(\mathsf{l}_i) \leq 2$; [Papadopoulo and Faugeras 1998]. Particularly, they show that the set $\mathcal{G}(\boldsymbol{\lambda}, \boldsymbol{\rho})$ of homogenous rank-2 matrices G for *given* left kernel $\boldsymbol{\lambda} \sim (\lambda_1, \lambda_2, \lambda_3)^\top$ and right kernel $\boldsymbol{\rho} \sim (\rho_1, \rho_2, \rho_3)^\top$ has 4 degrees of freedom. This can be seen by the relations $\boldsymbol{\lambda}^\top \mathsf{G} = \mathbf{0}^\top$ and $\mathsf{G}\boldsymbol{\rho} = \mathbf{0}$, each of which produces 3 constraints. However, since each of these two groups of constraints includes implicitly the singularity of G, only 5 of the resulting 6 equations are linearly independent. Therefore, only 4 of the 9 elements of any homogenous singular 3×3 matrix G with given left and right kernel are free.

Consequently any matrix of $\mathcal{G}(\boldsymbol{\lambda}, \boldsymbol{\rho})$ can be linearly expressed in terms of four rank-1 basis matrices G_1, G_2, G_3, G_4. There is no general way to define a basis, that is valid for all choices of $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$, and so in their derivations Papadopoulo and Faugeras assume the first coordinate of $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$ to have the largest absolute value. Then the following basis matrices are applicable:

$$\mathsf{G}_1 = \begin{bmatrix} \rho_3\lambda_3 & 0 & -\rho_1\lambda_3 \\ 0 & 0 & 0 \\ -\rho_3\lambda_1 & 0 & \rho_1\lambda_1 \end{bmatrix}, \quad \mathsf{G}_2 = \begin{bmatrix} -\rho_2\lambda_3 & \rho_1\lambda_3 & 0 \\ 0 & 0 & 0 \\ \rho_2\lambda_1 & -\rho_1\lambda_1 & 0 \end{bmatrix},$$
$$\mathsf{G}_3 = \begin{bmatrix} -\rho_3\lambda_2 & 0 & \rho_1\lambda_2 \\ \rho_3\lambda_1 & 0 & -\rho_1\lambda_1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathsf{G}_4 = \begin{bmatrix} \rho_2\lambda_2 & -\rho_1\lambda_2 & 0 \\ -\rho_2\lambda_1 & \rho_1\lambda_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This basis is valid as long as $\lambda_1 \neq 0$ and $\rho_1 \neq 0$, and for numerical reasons should be used only if these two elements have the largest absolute value. If other elements are the largest (there are in total 9 possible combinations), the corresponding basis can be found in a similar way. With respect to this basis we can express any matrix G, that has the left and right kernels $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$, as $\mathsf{G} = a_1\mathsf{G}_1 + a_2\mathsf{G}_2 + a_3\mathsf{G}_3 + a_4\mathsf{G}_4$ and as long as $a_2a_3 - a_1a_4 \neq 0$ the matrix G is of rank-2, and rank-1 otherwise.

Now we consider especially the correlation slices l_i of the trifocal tensor. Based on the fact that the left and right kernels of the three correlation slices are *concurrent* in the epipoles $\mathbf{e}_1''' \sim (e_1'''^{(1)}, e_1'''^{(2)}, e_1'''^{(3)})^\top$ and $\mathbf{e}_1'' \sim (e_1''^{(1)}, e_1''^{(2)}, e_1''^{(3)})^\top$ respectively, Papadopoulo and Faugeras show that the following relation must hold

$$e_1''^{(2)} e_1'''^{(2)} a_1^i + e_1''^{(2)} e_1'''^{(3)} a_2^i + e_1''^{(3)} e_1'''^{(2)} a_3^i + e_1''^{(3)} e_1'''^{(3)} a_4^i = 0, \tag{8.9}$$

where a_1^i to a_4^i are the coefficients needed to express l_i in the basis matrices of $\mathcal{G}(\boldsymbol{\lambda}_i, \boldsymbol{\rho}_i)$, with $i = 1, 2, 3$. And again, depending on the largest elements in $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$, different elements of the epipoles need to be selected. The form of relation (8.9) refers to the case where the first elements in $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$ have the largest absolute value.

With this preparations Papadopoulo and Faugeras show that the l_i matrices can be parameterized by 8 coefficients, provided the six nullspaces $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$ of the three correlation slices l_i are given. To see this, we consider the 12 coefficients a_i^j, $j = 1, \ldots, 4$, required for the linear expression of the three l_i matrices in the basis matrices of $\mathcal{G}(\boldsymbol{\lambda}_i, \boldsymbol{\rho}_i)$. Since for each i the four coefficients a_i^j satisfy the condition (8.9), we can express one of the a_i^j in the other quantities of (8.9). The best choice is to select that element a_i^j, whose coefficient in (8.9) hast the largest

absolute value. Furthermore, because the trifocal tensor is only defined up to scale we can drop one more of the remaining 9 coefficients a_i^j by setting it to 1.

Since the trifocal tensor has 18 degrees of freedom and 8 parameters are used for the parameterization of the three l_i matrices if the six nullspaces are given, 10 parameters remain for the parameterization of the six nullspaces $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$. The three left kernels $\boldsymbol{\lambda}_i$ are concurrent in the epipole $\mathbf{e}_1'''$, which is only defined up to scale and thus can be parameterized by two coordinates by normalizing the largest coordinate to 1. Then each of the three left kernels $\boldsymbol{\lambda}_i$ can be parameterized by an angle with respect to some basis vectors in the plane orthogonal to the direction of $\mathbf{e}_1'''$ and since the $\boldsymbol{\lambda}_i$ are also only defined up to scale, we can set $||\boldsymbol{\lambda}_i|| = 1$. So in total the set of concurrent left kernels $\boldsymbol{\lambda}_i$ is parameterized by 5 parameters.

The same reasoning can be applied to the three right kernels $\boldsymbol{\rho}_i$, which are concurrent in the epipole $\mathbf{e}_1''$. And so in total the set of concurrent right kernels $\boldsymbol{\rho}_i$ is also parameterized by 5 parameters.

This concludes the minimal parameterization introduced in [Papadopoulo and Faugeras 1998]. Note, however, that a single set of these 18 parameters can not represent all possible trifocal tensors, because the elements of the nullspaces $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$, the coefficients a_i^j and the coordinates of the epipoles must be mapped in a different way to the 18 minimal parameters depending on the configuration of the three images. The total number of maps is rather large: 3^6 for the selection of the element with the largest absolute value in the six nullspaces $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$ (this defines the basis matrices of $\mathcal{G}(\boldsymbol{\lambda}_i, \boldsymbol{\rho}_i)$); 9 for the selection of that coefficient a_i^j which is set to 1 to fix the scale ambiguity of the tensor; 3 for the selection of the largest coordinate of the epipole $\mathbf{e}_1''$ and analogously 3 for $\mathbf{e}_1'''$. This gives in total $3^6 \cdot 9 \cdot 3^2 = 59.049$ possible maps. The best map has to be selected judging from approximations obtained by e.g. the direct linear solution.

Prerequisite for the applicability of this minimal parameterization is that all six nullspaces $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$ are single lines; i.e. all three correlation slices l_i must have maximum rank-2. If $\operatorname{rank}(\mathsf{l}_i) \leq 2$ then the *nullspaces $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$ are not single lines* but pencils or the entire $\mathbb{P}^2$, and therefore the proposed method is not applicable. From table 7.4 on page 77 we see that this requires that neither the second, nor the third projection center coincides with the first one, as then all correlation slices would be of rank-1. In general this can be reached by adapting the roles of the three images.

However, there is still one drawback associated with this minimal parameterization, because we can see from table 7.4 that correlation slices l_i with $\operatorname{rank}(\mathsf{l}_i) \leq 2$ can occur for any image configuration; i.e. even if the three projection centers are distinct. Although this problem can be solved by changing the coordinate system of the first image in a way that all three correlation slices have rank-2, it shows that the applicability of this minimal parameterization depends on the coordinate system of the first image.

8.3.4 A new parameterization

In section 7.6.3 on page 99 we derived new sets of constraints for the trifocal tensor and proved that these sets are minimal. During the prove for the constraints obtained from the correlation slices we discovered, that we can parameterize the correlation slices and therefore the trifocal tensor without loss of generality in the form (7.57); i.e.

$$
\begin{aligned}
\mathsf{I}_1 &= \left[\, \mathbf{s}_1''', \quad v\mathbf{s}_1''' + m_1\mathbf{e}_1''', \quad w\mathbf{s}_1''' + n_1\mathbf{e}_1''' \,\right], \\
\mathsf{I}_2 &= \left[\, \mathbf{s}_2''', \quad v\mathbf{s}_2''' + m_2\mathbf{e}_1''', \quad w\mathbf{s}_2''' + n_2\mathbf{e}_1''' \,\right], \\
\mathsf{I}_3 &= \left[\, \mathbf{s}_3''', \quad v\mathbf{s}_3''' + m_3\mathbf{e}_1''', \quad w\mathbf{s}_3''' + n_3\mathbf{e}_1''' \,\right].
\end{aligned}
\tag{8.10}
$$

This parameterization, first presented in [Ressl 2002], is made up of 20 elements – the vectors $\mathbf{s}_i'''$, $\mathbf{m} = (m_1, m_2, m_3)^\top$, $\mathbf{n} = (n_1, n_2, n_3)^\top$ and the epipole $\mathbf{e}_1'''$ have 3 components each, plus v and w. From section 7.6.3 we know it is a consistent parameterization and therefore it has 18 degrees of freedom. Consequently the parameters have to satisfy 2 constraints, which simply correct two scaling ambiguities in 8.10. Obviously, the scale of $\mathbf{e}_1'''$ is undetermined as any change in that scale can be compensated in the parameters $\mathbf{m}$ and $\mathbf{n}$. Further, the overall scale of this parameterization represented in the scale of the vectors $\mathbf{s}_i'''$, $\mathbf{m}$ and $\mathbf{n}$ is also undetermined.

We can resolve these two scale ambiguities *explicitly*, by considering two suitable scaling constraints, e.g. $|\mathbf{e}_1'''| = 1$ and $||\,[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3''']\,||_{\text{Frob}} = 1$, or *implicitly* by adapting the parameters to form a *minimal parameterization*. First we consider the general, non-minimal, case and investigate the properties of this parameterization and afterwards we show how the minimal parameterization can be realized.

The non-minimal parameterization. The elements of the parameterization (8.10) have a simple geometric meaning; see proposition 8.1. Note, that in representation (8.10) there is one column in each correlation matrix I_i that is solely parameterized by the vector $\mathbf{s}_i'''$. We will refer to this unique column by the index c_I.

Proposition 8.1 (Geometric interpretation of the new parameterization.) *The elements in the parameterization* (8.10) *have the following geometric interpretation:*

- $\mathbf{e}_1'''$ *is the epipole of image* Ψ_1 *in image* Ψ_3.
- $(1, v, w)^\top$ *is the epipole* $\mathbf{e}_1''$ *of image* Ψ_1 *in image* Ψ_2 *– its component at position* c_I *is set to 1.*
- $[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3''']$ *is a (point) homography from image* Ψ_1 *to image* Ψ_3 *induced by the* c_I*th principal plane of image* Ψ_2*; i.e. the homography slice* J_{c_I}.
- $\mathbf{m}$ *is the 2-dimensional general eigenspace of* $\mathsf{J}_{c_I+1} - \kappa\mathsf{J}_{c_I}$*, which corresponds to the map of the* $(c_I - 1)$*st principal ray of image* Ψ_2 *into image* Ψ_1.
- $\mathbf{n}$ *is the 2-dimensional general eigenspace of* $\mathsf{J}_{c_I-1} - \kappa\mathsf{J}_{c_I}$*, which corresponds to the map of the* $(c_I + 1)$*st principal ray of image* Ψ_2 *into image* Ψ_1.

Proof:
The interpretation for $\mathbf{e}_1'''$ and $(1, v, w)^\top$ was already used during the derivation of the minimal set of constraints in section 7.6.3 on page 99. The interpretation of $[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3''']$ follows directly from table 7.10 on page 101; i.e. with $c_I = 1$ the J_j homographies are represented by this parameterization as

$$
\begin{aligned}
\mathsf{J}_1 &= \left[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3'''\right], \\
\mathsf{J}_2 &= v\mathsf{J}_1 + \mathbf{e}_1'''\mathbf{m}^\top, \\
\mathsf{J}_3 &= w\mathsf{J}_1 + \mathbf{e}_1'''\mathbf{n}^\top.
\end{aligned}
$$

What remains to show is that **m** and **n** represent the stated general eigenspaces. It is straight forward to show that $\kappa = v$ is a general eigenvalue with multiplicity 2 of $\mathsf{J}_{c_I+1} - \kappa\mathsf{J}_{c_I}$ and that $\kappa = w$ is a general eigenvalue with multiplicity 2 of $\mathsf{J}_{c_I-1} - \kappa\mathsf{J}_{c_I}$. From proposition 7.8 on page 73 about the general eigenvalue problem associated with the homography slices we see that the general eigenspace of the eigenvalue with multiplicity 2 is the map of the respective principal ray of image Ψ_2 into image Ψ_1.

□

The fact that in representation (8.10) the vectors $\mathbf{s}_i'''$ parameterize the same column, with index c_I, in all three I_i matrices is a possible limitation for the applicability of this parameterization. Only if the c_Ith column in each I_i matrix is different from $\mathbf{e}_1'''$ and the zero vector $\mathbf{0}$, the parameterization is applicable in general. Therefore and for numerical reasons, the index c_I is chosen so that the respective columns are farthest away from $\mathbf{e}_1'''$. Consequently this index c_I may be found by $\sum_{i=1}^{3} |\mathbf{e}_1''' \times \mathsf{I}_i \boldsymbol{i}_I| \rightarrow \max$; or we simply look for the largest absolute value of the coordinates of epipole $\mathbf{e}_1''$, because from proposition 8.1 we see that the c_Ith coordinate of this epipole is implicitly set to 1.

For the safe application of any parameterization it is important to know its critical configurations.

Proposition 8.2 (Critical configurations of the new parameterization.) *The parameterization* (8.10)*, taking into account the three possible maps for the index c_I, is applicable for any image configuration as long as not all three projection centers coincide.*

Proof:
Since it is part of the parameterization, that the length of $\mathbf{e}_1'''$ and one component in $\mathbf{e}_1''$ are set to 1, problems surely arise if either of these epipoles is the zero-vector. This would happen if the first projection center $\boldsymbol{Z}_1$ coincides with one of the other two; i.e. $\boldsymbol{Z}_1 = \boldsymbol{Z}_3$ or $\boldsymbol{Z}_1 = \boldsymbol{Z}_2$. This problem can be solved - as long as not all three projection centers coincide - by changing the role of the images in the way that the image with the unique projection center plays the role of image Ψ_1. In case of $\boldsymbol{Z}_1 = \boldsymbol{Z}_2 = \boldsymbol{Z}_3$ the respective trifocal tensor becomes the zero-tensor, which does not correspond to a valid relative orientation of three images.

Another problem with this parameterization could come from the fact, that the vectors $\{\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3'''\}$ parameterize the same column (with index c_I) in all three correlation slices. Still keep in mind that we choose the best column for this parameterization – the one that is farthest away from $\mathbf{e}_1'''$. If we take the minimal parameterization exactly as it is given in representation (8.10), we see, that the columns of $\mathsf{I}_1 = [\mathbf{a}_1''', \mathbf{a}_2''', \mathbf{a}_3''']$ are parameterized by $\mathbf{e}_1'''$ and the vector $\mathbf{s}_1'''$. Thus it must be assured, that $\mathbf{s}_1'''$ is different from $\mathbf{e}_1'''$ and different from the zero-vector, i.e. $\mathbf{s}_1''' \not\sim \mathbf{e}_1'''$, because otherwise the column vectors $\mathbf{a}_2''' \not\sim \mathbf{e}_1'''$ and/or $\mathbf{a}_3''' \not\sim \mathbf{e}_1'''$ can not be parameterized by $\mathbf{s}_1'''$ and $\mathbf{e}_1'''$. Of course, if $\mathbf{a}_2''' \sim \mathbf{e}_1'''$ *and* $\mathbf{a}_3''' \sim \mathbf{e}_1'''$, than we would have no problem. So, we will prove, that the first case, $\mathbf{s}_1''' \sim \mathbf{e}_1'''$ but $\mathbf{a}_2''' \not\sim \mathbf{e}_1'''$ and/or $\mathbf{a}_3''' \not\sim \mathbf{e}_1'''$, can not occur. This prove is outlined in the following.

First we arrange the three matrices I_i as the rows of a large matrix U, which then has 9 rows and 3 columns. We, however, consider the elements of U being the column vectors of the I_i matrices. So, U has 3×3 elements and the element at row ξ and column η is the ηth column vector in matrix I_ξ, which is the image of the intersection point of principal plane $\boldsymbol{\Pi}_\eta''$ with the principal ray $\boldsymbol{R}_\xi'$.

Now, we assume that one element (row ξ, column η) of U is $\sim \mathbf{e}_1'''$. This may happen due to two situations: A1) $\boldsymbol{Z}_3 \in \boldsymbol{R}'_\xi$ or A2) $\boldsymbol{Z}_1 \in \boldsymbol{\Pi}''_\eta$. Then, we consider that one element (row ξ, column η) of U shall be $= \mathbf{0}$. This may happen also due to two situations: B1) $\boldsymbol{Z}_3 = \boldsymbol{\Pi}''_\eta \wedge \boldsymbol{R}'_\xi$ or B2) $\boldsymbol{R}'_\xi \in \boldsymbol{\Pi}''_\eta$. However, B1) implies $\boldsymbol{Z}_3 \in \boldsymbol{R}'_\xi$ ($\rightarrow$ A1) and B2) implies $\boldsymbol{Z}_1 \in \boldsymbol{\Pi}''_\eta$ ($\rightarrow$ A2). Thus, the only possible situations, that may return one element in U being $\sim \mathbf{e}_1'''$ or $= \mathbf{0}$ are the ones of A1) and A2).

These situations, however, not only return the element at row ξ and column η of matrix U to be $\sim \mathbf{e}_1'''$ or $= \mathbf{0}$, they further imply: A1) returns that all elements in row ξ of U are $\sim \mathbf{e}_1'''$; i.e. the entire matrix I_ξ, which then has rank-1. And so all columns of I_ξ can be parameterized by $\mathbf{s}_\xi$, being $\sim \mathbf{e}_1'''$ or $= \mathbf{0}$, and $\mathbf{e}_1'''$. A2) returns that all elements in column η of U are $\sim \mathbf{e}_1'''$; i.e. the column η in all three correlation matrices I_i. Again, the parameterization of these columns is not difficult, but what is more important: When situation A2) occurs, the ηth column in the three correlation matrices I_i will never be used as the vectors $\{\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3'''\}$ in the parameterization, since they are not far away from $\mathbf{e}_1'''$.

This completes the prove, that it is impossible, that one of the three vectors $\{\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3'''\}$ is $\sim \mathbf{e}_1'''$ or $= \mathbf{0}$, but one of the other columns in I_i is different from $\mathbf{e}_1'''$ and $\mathbf{0}$. Thus, the new parameterization of the trifocal tensor holds for any image configuration – provided not all three projection centers coincide.

□

If the parameterization (8.10) is to be used in the Gauß-Helmert model, then we have to make the same adaption as for the parameterization using the projection matrices in section 8.3.1; i.e. the unknowns are not the 27 tensor elements which make up the vector $\boldsymbol{t}$ introduced in equation (8.2) on page 112, but the 20 elements of the new parameterization, which make up the vector $\boldsymbol{q}$. Consequently the Jacobian matrix A of the derivatives of the trilinearities with respect to the unknowns must be adapted also. We can represent the vector $\boldsymbol{q}$ as

$$\boldsymbol{q} = \left(\mathbf{s}_1'''^\top, \mathbf{s}_2'''^\top, \mathbf{s}_3'''^\top, v, \mathbf{m}^\top, w, \mathbf{n}^\top, \mathbf{e}_1'''^\top\right)^\top. \tag{8.11}$$

With the vectors $\boldsymbol{t}$ and $\boldsymbol{q}$ we can represent the derivative of any trilinearity $T(\cdot,\cdot,\cdot)$ as

$$\frac{\partial T(\cdot,\cdot,\cdot)}{\partial \boldsymbol{q}} = \frac{\partial T(\cdot,\cdot,\cdot)}{\partial \boldsymbol{t}} \frac{\partial \boldsymbol{t}}{\partial \boldsymbol{q}}.$$

If we further represent the respective Jacobian matrices by A_t and A_q, where A_t follows from the trilinear derivatives of table 8.1 on page 112, then A_q is easily obtained as

$$\mathsf{A}_q = \mathsf{A}_t \mathsf{D}_q, \tag{8.12}$$

with

$$\underset{27\times 20}{\mathsf{D}_q} = \left[\begin{array}{c|cc|cc|c} \mathsf{I} & & & & & \\ v\mathsf{I} & \mathbf{s}_1''' & \mathbf{e}_1''' & & & m_1\mathsf{I} \\ w\mathsf{I} & & & \mathbf{s}_1''' & \mathbf{e}_1''' & n_1\mathsf{I} \\ \hline \mathsf{I} & & & & & \\ v\mathsf{I} & \mathbf{s}_2''' & \mathbf{e}_1''' & & & m_2\mathsf{I} \\ w\mathsf{I} & & & \mathbf{s}_2''' & \mathbf{e}_1''' & n_2\mathsf{I} \\ \hline \mathsf{I} & & & & & \\ v\mathsf{I} & \mathbf{s}_3''' & \mathbf{e}_1''' & & & m_3\mathsf{I} \\ w\mathsf{I} & & & \mathbf{s}_3''' & \mathbf{e}_1''' & n_3\mathsf{I} \end{array}\right], \tag{8.13}$$

where the positions of zero elements have been left blank for clearer representation. The identity matrix of dimension 3 is represented as I.

For the usage of this parameterization in the Gauß-Helmert model we further need approximate values for the elements of this parameterization, which can be obtained from the direct linear solution following table 8.3.

(i) Compute the epipole $\mathbf{e}_1'''$ from the direct linear solution using the method of Spetsakis and Aloimonos in section 7.4.1 on page 83. This also gives $|\mathbf{e}_1'''| = 1$.

(ii) Choose the index c_I as explained above; this gives the vectors $\mathbf{s}_1'''$, $\mathbf{s}_2'''$ and $\mathbf{s}_3'''$, which are the columns of the matrix J_{c_I}. Divide those three vectors and the trifocal tensor from the direct linear solution by $||\mathsf{J}_{c_I}||_{\text{Frob}}$. This ensures the second scaling constraint $||\,[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3''']\,||_{\text{Frob}} = 1$ is satisfied.

(iii) Find the other elements $(\mathbf{m}, \mathbf{n}, v, w)$ of the parameterization by solving the following overdetermined system, which is linear in these missing parameters,

$$\underbrace{\begin{bmatrix} \mathbf{s}_1''' & \mathbf{e}_1''' & & \\ \mathbf{s}_2''' & & \mathbf{e}_1''' & \\ \mathbf{s}_3''' & & & \mathbf{e}_1''' \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} v & w \\ m_1 & n_1 \\ m_2 & n_2 \\ m_3 & n_3 \end{bmatrix}}_{V} = \underbrace{\begin{bmatrix} \text{vec}\left(\mathsf{J}_{c_I+1}\right), & \text{vec}\left(\mathsf{J}_{c_I-1}\right) \end{bmatrix}}_{N}$$

by minimizing the squared distance to the $(c_I + 1)$st and the $(c_I - 1)$st column of the I_i matrices; $V = (M^\top M)^{-1} M^\top N$.

(iv) (optional) Refine this parameterization iteratively by solving the following least squares problem:

$$|D_q \Delta\boldsymbol{q} - (\boldsymbol{t} - \boldsymbol{t}(\boldsymbol{q}))|^2 \quad \rightarrow \quad \min.$$

The vector $\boldsymbol{t}$ represents the direct linear solution of the trifocal tensor, $\boldsymbol{t}(\boldsymbol{q})$ is the valid trifocal tensor computed from the parameterization $\boldsymbol{q}$ using relation (8.10), $\Delta\boldsymbol{q}$ is the correction for the elements of the parameterization $\boldsymbol{q}$ and D_q is defined in (8.13). Usually 1-3 iterations are sufficient.

Table 8.3: *Initialization of the new parameterization* (8.10) *using the direct linear solution for the trifocal tensor.*

Note that this way of getting the approximation for the new parameterization (including the optional last step) corresponds to a constrained solution of type (Bb), or *projection method*, mentioned at the beginning of this chapter; i.e. the input for the constrained solution is the direct linear solution of the tensor and using the new parameterization that valid trifocal tensor is found, which lies closest to the direct linear solution.

Observe that the new parameterization and the parameterization using the projection matrices are closely related. The latter turns into the new parameterization, if the six constraints for the projection matrices are adapted as shown in table 8.4.

ambiguity	parameterization using the projection matrices	new parameterization
fix scale of $\mathsf{P}_2 = [\mathsf{A}, \mathbf{e}_1'']$	$\lvert\mathbf{e}_1''\rvert = 1$	$\max(\lvert e_1''\rvert^{(j)}) = \lvert e_1''^{(c_I)}\rvert = 1$
fix scale of $\mathsf{P}_3 = [\mathsf{B}, \mathbf{e}_1''']$	$\lvert\mathbf{e}_1'''\rvert = 1$	$\lvert\mathbf{e}_1'''\rvert = 1$
fix $\mathbf{v}$	$\mathbf{a}_i^\top \mathbf{e}_1'' = 0, i = 1,2,3$	$\mathbf{a}_{c_I} = \mathbf{0}$
fix w	$\lvert\lvert\mathsf{A}\rvert\rvert_{\text{Frob}} = 1$	$\lvert\lvert\mathsf{J}_{c_I}\rvert\rvert_{\text{Frob}} = 1$

Table 8.4: *Comparison of the new parameterization and the parameterization using the projection matrices (section 8.3.1 on page 117) with respect to the choices of resolving the 6 ambiguities in the 24 elements of the projection matrices P_2 and P_3. The vector $\mathbf{v}$ and the scalar w are the remaining degrees of freedom in the projective transformation that takes P_1 to the standard form $[I_3, \mathbf{0}]$; see section 7.4.3 on page 90. The vector $\mathbf{a}_i$ is the ith column of the matrix A.*

The minimal parameterization In the general form of the new parameterization we have to consider two scaling constraints during the computation. We can avoid these constraints by adapting the parameterization in the following simple way. We can set the largest absolute value of $\mathbf{e}_1'''$ and $[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3''']$ to 1. Consequently the whole parameterization depends only on 18 parameters and so we derived a minimal parameterization for the trifocal tensor in very simple way.

The elements of this minimal parameterization can be initialized in the same way as the non-minimal one using table 8.3 and afterwards setting the largest absolute value in $\mathbf{e}_1'''$ and $[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3''']$ to 1.

Considering that we have 3 choices for selecting the index c_I of the columns, which are parameterized by $\mathbf{s}_i'''$, 3 choices for the largest absolute value of $\mathbf{e}_1'''$ and 9 choices for the largest absolute value of $[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3''']$, this minimal parameterization has in total 81 maps, which is a number much smaller than for the minimal parameterization of Papadopoulo and Faugeras in section 8.3.3.

Consequently the implementation of this minimal parameterization will be much simpler than the one of section 8.3.3. The implementation of the non-minimal parameterization together with its two constraints, however, is still easier, because only the 3 choices for selecting the index c_I need to be considered. Therefore we will consider in the following only the non-minimal parameterization.

8.3.5 Considering the internal constraints in the Gauß-Helmert model

For estimating a valid trifocal tensor by minimizing reprojection error we can use the Gauss-Helmert model introduced in section 8.2. Now we have to decide how to realize the constraints in this model in an easy way. We can either estimate the trifocal tensor with its 27 elements and introduce suitable 8 constraints from section 7.6 plus the scaling constraint to achieve 18 degrees of freedom, or we can use one of the parameterizations presented in the previous sections, which, being minimal or non-minimal, have always 18 degrees of freedom.

The *set of constraints* introduced in section 7.6 are not the first candidates because the following restrictions are associated to them.

The non-minimal set of Papadopoulo and Faugeras (section 7.6.1 on page 96)**:** The Gauß-Helmert model does not offer the possibility of including redundant constraints[1]. Therefore suitable 8 constraints must be selected from the 12 constraints proposed by Papadopoulo and Faugeras, and it is not obvious which ones are the best for this purpose. Further, the epipolar constraints (7.48) on page 96, which must be included in any case, require that all three correlation slices have maximum rank-2. The rank of these matrices depends on the orientation of the measurement system $\{x', y'\}$ of the first image. So by proper choice of the system $\{x', y'\}$ the maximum rank case can be established in general. However, it may not be obvious in advance how to choose the coordinate system in image Ψ_1 to provide rank-2 for all correlation slices. So, due to this dependency on $\{x', y'\}$ the usage of these constraints is further reduced.

The minimal set of Canterakis (section 7.6.2 on page 97)**:** Although this is a minimal set, the realization of these constraints in the Gauß-Helmert model seems to be a difficult task, due to the high complexity of the constraints. For example, the implementation of the two constraints for demanding identical (up to scale) eigenvectors for the single general eigenvalues, requires the differentials of these two constraints with respect to the 27 tensor elements, which appears to be complicated even by numerical differentiation. Therefore the set of Canterakis does not seem to be the first choice for the Gauß-Helmert model.

The new minimal sets of constraints (section 7.6.3 on page 99)**:** The new sets include either the epipolar constraint (7.53) on page 101 or the concurrency constraint (7.59) on page 105 and for both the same restriction as for the set of Papadopoulo and Faugeras holds. Only if all correlation slices have maximum rank-2 (for the epipolar constraint) the first set of constraints can be applied; and only if all homography slices have at least rank-2 the second set of constraints can be applied. So, although these sets are minimal, they are also not the first choice for the Gauß-Helmert model. For the set derived from the correlation slices some experiments were made and it became evident, that if not all correlation slices have maximum rank-2, then the iteration does not converge, but oscillates around the actual solution.

One drawback that holds for all previous sets of constraints, is that the constraints expressed in the 27 tensor elements are of high order, ranging from degree 3 to degree 12.

The other possibility of introducing the constraints in the Gauß-Helmert model is to use one of the *parameterizations with 18 degrees of freedom* presented in the last sections. Concerning their implementation in the Gauß-Helmert model we can state the following.

The parameterization using the projection matrices (section 8.3.1 on page 117)**:** This is a simple bilinear parameterization using 24 parameters. It is is independent on the choice of the coordinate systems of the three images and valid for all image configurations, as long as not all three projection centers coincide and the projection center of the first image is unique. To realize the 18 degrees of freedom in the Gauß-Helmert model we need to consider 6 constraints, which were discussed in section 8.3.1. The parameterization itself and the 6 necessary constraints have a simple form. Consequently their implementation is easy and so this parameterization is a perfect candidate for the Gauß-Helmert model.

[1] A workaround would be to introduce the constraints as conditions with very large weights to assure they are satisfied. This, however, may run the risk that numerical problems arise or that the internal constraints are not fully satisfied.

The six-point parameterization (section 8.3.2 on page 118)**:** Although this parameterization allows a *perfect method for detecting gross errors in the data*, it is not the first choice for parameterizing the trifocal tensor in the Gauß-Helmert model. Again we are confronted with a selecting problem that can not be solved in advance: Which points of all n point correspondences are to be used for the six-point parameterization and which coordinates of those are to be kept fix and which allow to be free? This corresponds to $\binom{n}{6} \cdot 18$ choices. Further we need the matrix $\mathbf{A}_{\mathbf{r}}$, containing the differentials of the trilinearities with respect to the 18 free coordinates of the the six point parameterization, which make up the vector $\mathbf{r}$. This can be done analogously to section 8.3.1 as $\mathbf{A}_{\mathbf{r}} = \mathbf{A}_{\mathbf{t}}\mathbf{D}_{\mathbf{r}}$. The elements of $\mathbf{D}_{\mathbf{r}}$ contain the differentials of the 27 tensor elements with respect to the 18 free coordinates of the the six point parameterization. The determination of the elements of $\mathbf{D}_{\mathbf{r}}$ is not trivial and seems to be realizable only by numerical differentiation, which itself is rather involving: First the six points must be transformed to their canonical representation, then the cubic equation for λ_1 must be solved, which gives then up to 3 solutions for the projection matrices, the appropriate set of projection matrices is the transformed back to the system of the original points. Finally these projection matrices determine the 27 elements of the tensor.

So, all in all, the realization of the six-point parameterization in the Gauß-Helmert model would be rather involving.

The minimal parameterization by Papadopoulo and Faugeras (section 8.3.3 on page 122)**:** The main drawbacks associated with this parameterization are the following. It is rather involving to setup this minimal parameterization at the beginning of the computation, as there are several selections to be made concerning the fixing and substitution of elements, resulting in a total number of 59.049 possible maps. During the iteration it may turn out that other elements are better suited for fixing or substitution, therefore these criteria need to be observed throughout the iterations and if needed the parameterization needs to be adapted. Since there are several cases to be considered, the respective program code will be rather long and complex. And since this parameterization requires all correlation slices to have maximum rank-2, another drawback of this parameterization is its dependency on the choice of the coordinate system of the first image. Although this does not practically hamper the realization of this parameterization, it is nevertheless an unnecessary additional regard.

The new parameterization (section 8.3.4 on page 123)**:** The new parameterization can be realized in a minimal or non-minimal way. The minimal form has 81 possible maps, therefore we will consider the non-minimal parameterization, as it is easier to realize. This new parameterization offers a simple bilinear relation between the 27 tensor elements and the 20 elements of the parameterization. It is is independent on the choice of the coordinate systems of the three images and valid for all image configurations, as long as not all three projection centers coincide and the projection center of the first image is unique.

Due to its simplicity and its general applicability, this new parameterization is suited for the realization of the 18 degrees of freedom in the Gauß-Helmert model.

Summing it up, of all the sets of constraints and parameterizations presented, only the parameterization using the projection matrices and the new parameterization are suited for computing a valid trifocal tensor in the Gauß-Helmert model. Both are *non-minimal* parameterizations. The

implementation of any minimal parameterizations is not recommended, as the program code will be relatively complex due to the many case differentiations involved.

The benefit of the parameterization using the projection matrices is that there is only one map from the trifocal tensor to this parameterization, whereas for the new parameterization there are three maps, i.e. the choice of the column index c_I. On the other hand the parameterization using the projection matrices requires 6 constraints, whereas the new parameterization needs only 2 constraints.

However, since the simplicity of the implementation, i.e. the number of maps, is more important than the number of required constraints, the *parameterization using the projection matrices is the recommended way to represent a valid trifocal tensor in the Gauß-Helmert model*.

8.4 Transformation of the trifocal tensor due to conditioning

As it was already mentioned in section 3.7, it is inevitable to do a conditioning if measurements are introduced as projective quantities in the computations. The effect of not conditioning the measurements can be exemplified especially by the point trilinearities (the last row in table 8.1 on page 112). From there we see that the elements of the sub-matrix A_c are products of three *homogenous* point coordinates, one from each image. The homogenous part x_H is 1 for all real points, whereas the coordinates of the Euclidian part $\boldsymbol{x}_O$ can have any value. Suppose the points are measured in digital images with 2000×3000 pixels, the Euclidian part will be in the order of 1000. Consequently the elements of A_c will vary between 1^3 and 1000^3, and we will encounter numerical problems in the solution of 8.1.

With the conditioning matrices presented in section 3.7 on page 28 we can easily solve this problem. However, in this case the computed tensor, represented by the vector $\bar{\mathbf{t}}$, refers to the system of the conditioned measurements. The transformation back to the system of the original measurements is done in the following way.

If the points in each image are conditioned by different matrices, i.e. $\bar{\mathbf{x}}' \sim \mathsf{H}'\mathbf{x}'$, $\bar{\mathbf{x}}'' \sim \mathsf{H}''\mathbf{x}''$ and $\bar{\mathbf{x}}''' \sim \mathsf{H}'''\mathbf{x}'''$, where $\mathbf{x}$ represents the original measurement and $\bar{\mathbf{x}}$ the conditioned one. Then the correlation slice I_i of the original tensor are computed from the conditioned ones $\bar{\mathsf{I}}_i$ by

$$\begin{aligned}
\mathsf{I}_1 &= m'\mathsf{H}'''^{-1}\bar{\mathsf{I}}_1\mathsf{H}''^{-\top} \\
\mathsf{I}_2 &= m'\mathsf{H}'''^{-1}\bar{\mathsf{I}}_2\mathsf{H}''^{-\top} \\
\mathsf{I}_3 &= \mathsf{H}'''^{-1}(-m'x'_C\bar{\mathsf{I}}_1 - m'y'_C\bar{\mathsf{I}}_2 + \bar{\mathsf{I}}_3)\mathsf{H}''^{-\top},
\end{aligned} \tag{8.14}$$

where m', x'_C and y'_C are the elements of H'; cf. equation (3.49) on page 29. This transformation follows simply from the point trilinearities (see table 7.5 on page 80) by applying the three conditioning matrices $\mathsf{H}^{()}$ and the axiator relation (B.12) on page 179.

This transformation can also be performed in the following way. Let the original tensor be represented by the vector $\mathbf{t}$ and the conditioned tensor by the vector $\bar{\mathbf{t}}$. Then it holds:

$$\mathbf{t} \sim \mathsf{H}_T\bar{\mathbf{t}}, \tag{8.15}$$

where

$$\mathsf{H}_T \sim \mathsf{H}'^{\top} \otimes \mathsf{H}''^{-1} \otimes \mathsf{H}'''^{-1}.$$

Although the version (8.15) requires much more computation time than the version (8.14), it is particularly useful, because it enables the transformation of the covariance matrix $\mathbf{\Sigma}_{\bar{t}\bar{t}}$ of the conditioned tensor $\bar{\mathbf{t}}$ to the covariance matrix $\mathbf{\Sigma}_{tt}$ of the original tensor $\mathbf{t}$ by

$$\mathbf{\Sigma}_{tt} = \mathsf{H}_T \mathbf{\Sigma}_{\bar{t}\bar{t}} \mathsf{H}_T^\top.$$

8.5 Critical configurations for determining the trifocal tensor

When computing the trifocal tensor we also have to take care of critical configurations or surfaces, which do not allow a unique solution for the tensor. Due to [Hartley 2000] the following configurations are critical when determining the trifocal tensor from *image point correspondences*:

- All image point correspondences arise from object points lying in the same *plane*, which does not contain the three projection centers [2].
- All image point correspondences arise from object points lying on a certain space curve of degree 3, called *twisted cubic*, which also contains one of the three projection centers.
- All image point correspondences arise from object points lying in a six-parameter family of *space curves of degree 4*, which also contain the three projection centers.

Again as in the case of the fundamental matrix, point correspondences arising from the same plane are critical, which is not very surprising, as in this case the points between any pair of images are related by a homography. This is also understandable from the photogrammetric point of view, because the three images are considered *uncalibrated*. Consequently a calibration is associated with the determination of the trifocal tensor, which is well known to be impossible for planar point correspondences. However, what is different from the critical surfaces of the fundamental matrix (which are ruled quadrics) is, that for the trifocal tensor there are no critical *surfaces*, but only critical *curves* of certain degree. This is of practical importance, as the chance of running into a critical situation is much smaller when computing the trifocal tensor compared to computing the fundamental matrix (from the same points).

In contrast to the fundamental matrix, the determination of the trifocal tensor can also be done using image line correspondences. Here critical configurations also exist, with the so-called *linear line complex* being one of practical relevance; [Stein and Shashua 1998]. A linear line complex (LLC) is a set of 3D lines intersecting on a common line; e.g. all lines that are visible on two facades of a house form a LLC because they intersect on the edge of the two facades (see figure 8.1). In the case of this critical configuration the system for the direct linear solution (8.1) on page 112 has rank-23 instead of rank-26. A *unique* solution can be found nevertheless by imposing the internal constraints of the trifocal tensor. More on the topic of critical line configurations can be found in [Buchanan 1992].

[2] The rank of the matrix A in system (8.1) has rank-21 instead of rank-26; [Stein and Shashua 1998]. Therefore the right nullspace is a 6-dimensional subspace in $\mathbb{R}^{27}$. So by intersecting this nullspace with the manifold of dimension 18 defined by the internal constraints of the tensor (i.e. by looking for the members in the nullspace that satisfy the constraints) a valid trifocal tensors could be obtained. This intersection, however, will not be unique in general.

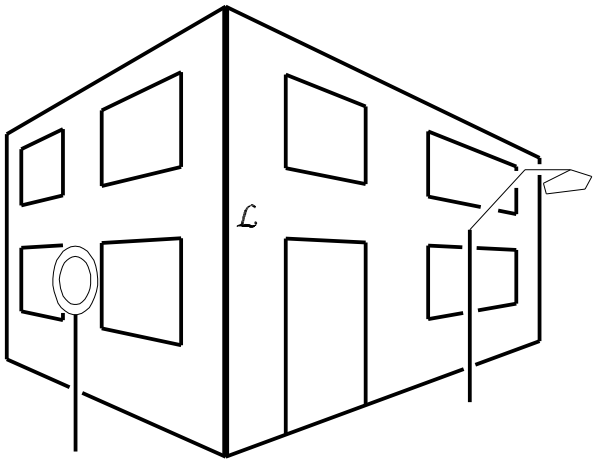

Figure 8.1: *A critical configuration of lines. All thick lines intersect the line $\mathcal{L}$ and form a linear line complex.*

Chapter 9

Empirical Investigations concerning the Quality of the estimated Trifocal Tensor

In this chapter we will compute the trifocal tensor by different methods for different image configurations and for a varying number of corresponding points. These examples are based on synthetic data and shall demonstrate

- the differences between minimizing algebraic and reprojection error,
- the effects of considering or neglecting the internal constraints of the tensor, and
- the impact of critical configurations; i.e. how close must points lie to the same plane so that the computation fails? To answer this question the object points will be placed inside a cuboid, which is then incrementally compressed in one direction till the computation fails; the compression for which the computation is still possible will be referred to as **minimum thickness** of the cuboid.

The examples are conducted and evaluated in the following way.

(i) Define a configuration of three images by means of exterior and interior orientation. This gives the original projection matrices P_i.

(ii) Define a cuboid in object space, place n points $\mathbf{X}_j$ regularly inside and project those points into the three images.

(iii) Add normally distributed noise (e.g. with 1 pixel standard deviation) to those image points and randomly select $m < n$ point correspondences.

(iv) Determine the trifocal tensor from these point correspondences by some selected method, which minimizes algebraic or reprojection error with or without considering the internal constraints.

(v) Retrieve projection matrices $\overline{\mathsf{P}}_i$ from the computed tensor using the method of section 7.4.3 on page 90.

(vi) Determine all n object points $\overline{\mathbf{X}}_j$ by spatial intersection (minimizing reprojection error) using the projection matrices $\overline{\mathsf{P}}_i$.

(vii) Since those retrieved $\overline{\mathsf{P}}_i$ differ from the original P_i by some unknown 3D projective transformation H^{-1} as $\mathsf{P}_i \sim \overline{\mathsf{P}}_i\mathsf{H}^{-1}$, determine H using the known original points $\mathbf{X}_j$ and the intersected points $\overline{\mathbf{X}}_j$ by minimizing $\sum d_j^2$, with $d_j = |\boldsymbol{X}_j - \boldsymbol{Y}_j|$ and $\mathbf{Y}_j \sim \mathsf{H}\overline{\mathbf{X}}_j$. Do not use the m select points which were used to compute the trifocal tensor.

(viii) Store the mean and maximum of the d_j values. If the mean value is above a certain threshold, then the respective solution is classified as 'bad'.

(ix) Repeat steps (iv) – (viii) for different methods of computing the trifocal tensor by minimizing algebraic and reprojection error with and without considering the internal constraints.

(x) Repeat steps (iii) – (ix) 1000 times.

(xi) Repeat steps (ii) – (x) by incrementally compressing the cuboid till all object points $\mathbf{X}_j$ lie in one plane.

(xii) Plot the averages over all 1000 samples of the mean and maximum d_j values (referred to as **ground errors**), and the percentage of 'bad' solutions against the compression rate of the cuboid.

(xiii) Repeat steps (ii) – (xii) for different m.

9.1 The setup

In this section the setup of the examples is described; i.e. the image configurations and the computation methods.

Image configurations. Five different image configurations 'Tetra', 'Air1', 'Air2', 'Street1' and 'Street2' have been created.

'Tetra': The three images and the center of the object points form a tetrahedron. The camera geometry is based on the Nikon DCS 460 camera. So each image has 3000×2000 pixels, with a pixel sice of 9×9 µm^2, therefore the whole CCD chip measures 27×18 mm^2. The principal point lies at the center of the image at $(1500, -1000)$, the principal distance c is given by 3500 pixel, which corresponds to about 32 mm. No affinity is assumed for the interior orientation.

The object points are regularly placed inside a cuboid of dimension $3\times1.5\times3$ m^3, with 8 points along each direction. All three principal axes intersect in the center of the cube with an angle of about 90°. To find the critical planarity situation this cube will be compressed in the direction orthogonal to the trifocal plane symmetrically to the center of the cuboid.

The projection matrices of the three images are given in table 9.1 on page 138 and the whole configuration is depicted in figure 9.1 on page 138.

'Air1': The images are assumed to be three consecutive images of an aerial flight strip flown with 80% length overlap. The projection centers lie 1500 m above the center of the object points, which is placed at height 0. The camera geometry is assumed to be the *normal angle* case. The image size is 23×23 cm^2 and the images are scanned with 15 µm resulting

therefore in 15300×15300 pixels per image. The principal point is placed in the center of the image at $(0, 0)$ (i.e. the image measurements refer to the camera system due to a fiducial transformation). The principal distance is assumed to be 300 mm, which corresponds to 20000 pixel. Again no affinity is assumed for the interior orientation.

The object points are regularly placed inside a cuboid of dimension $690 \times 1150 \times 225$ m^3, with 8 points along each direction. The corresponding projection rays intersect in the center of the cube with an angle of about 17°. To find the critical planarity situation this cuboid will be compressed in the vertical direction symmetrically to the center of the cuboid. In the beginning the height of the cuboid is 15% of the flight height.

The projection matrices of the three images are given in table 9.2 on page 139 and the whole configuration is depicted in figure 9.2 on page 139.

'Air2': This configuration is identical to 'Air1', except that the third image is assumed to be from the neighboring strip, which overlaps with the first strip by 60%. The object points are regularly placed inside a cuboid of dimension $920 \times 690 \times 225$ m^3, with 8 points along each direction. The corresponding projection rays intersect in the center of the cube with an angle of about 19°. To find the critical planarity situation this cuboid will be compressed in the vertical direction symmetrically to the center of the cuboid. In the beginning the height of the cuboid is 15% of the flight height.

The projection matrices of the three images are given in table 9.3 on page 140 and the whole configuration is depicted in figure 9.3 on page 140.

'Street1': The images are assumed to be three consecutive images acquired from a car driving on a straight and horizontal street approaching a vertical object; e.g. a garage. The image geometry is identical to the 'Tetra' configuration. The distance between the vertical object and its closest projection center is 10 m, which is also the depth of the object. The front facade measures 8.6×5.7m^2. This facade is mapped in the first image by covering 100% of the image size, and consecutive images overlap by 80% in area. The object points are regularly placed inside this vertical object, which is represented as a cuboid of dimension $8.6 \times 5.7 \times 10$ m^3, with 8 points along each direction. The corresponding projection rays intersect in any point of the cubic with an angle of less than 3°. To find the critical planarity situation this cuboid will be compressed in the direction of the street towards the front facade.

The projection matrices of the three images are given in table 9.4 on page 141 and the whole configuration is depicted in figure 9.4 on page 141.

'Street2': This configuration is identical to 'Street1', except for finding the critical planarity situation. Here the cuboid will be compressed in vertical direction towards the street; i.e. how close to the street must the points lie for the computation to fail? For this purpose the cuboid is adapted to the dimension $8.6 \times 2.85 \times 10$ m^3.

The projection matrices of the three images are given in table 9.5 on page 142 and the whole configuration is depicted in figure 9.5 on page 142.

$$
\mathsf{P}_1 = \begin{bmatrix} -0.929723 & -0.176307 & -0.536846 & 0.000000 \\ -0.117883 & 1.029001 & 0.094305 & 0.000000 \\ -0.000203 & -0.000118 & 0.000163 & 0.000000 \end{bmatrix}
$$

$$
\mathsf{P}_2 = \begin{bmatrix} -0.319666 & -0.176306 & 1.024892 & 3.196663 \\ 0.117883 & 1.029001 & 0.094305 & -1.178830 \\ 0.000203 & -0.000118 & 0.000163 & -0.002034 \end{bmatrix}
$$

$$
\mathsf{P}_3 = \begin{bmatrix} 1.000000 & 0.352214 & 0.244169 & -8.050175 \\ 0.000000 & -0.804538 & 0.659054 & 6.967296 \\ 0.000000 & 0.000235 & 0.000163 & -0.002033 \end{bmatrix}
$$

Table 9.1: *The projection matrices of configuration 'Tetra'.*

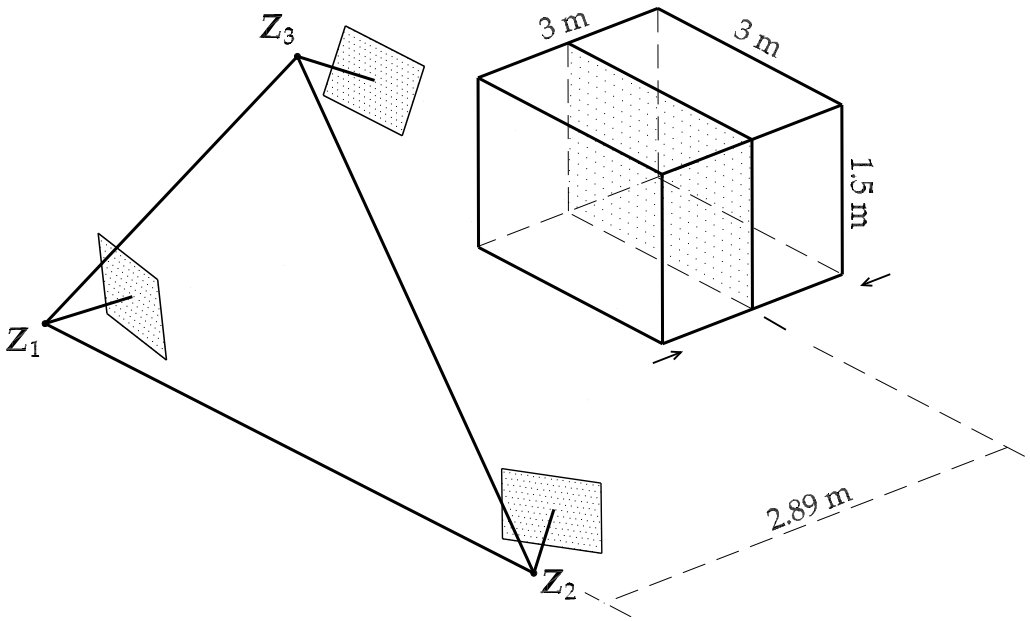

Figure 9.1: *Sketch of configuration 'Tetra'. The arrows point in direction of the cuboid compression, the limit of which is represented by the small horizontal bar.*

$$
\mathsf{P}_1 = \begin{bmatrix} 1.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & -0.000050 & 0.075000 \end{bmatrix}
$$

$$
\mathsf{P}_2 = \begin{bmatrix} 1.000000 & 0.000000 & 0.000000 & -230.000000 \\ 0.000000 & 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & -0.000050 & 0.075000 \end{bmatrix}
$$

$$
\mathsf{P}_3 = \begin{bmatrix} 1.000000 & 0.000000 & 0.000000 & -460.000000 \\ 0.000000 & 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & -0.000050 & 0.075000 \end{bmatrix}
$$

Table 9.2: *The projection matrices of configuration 'Air1'.*

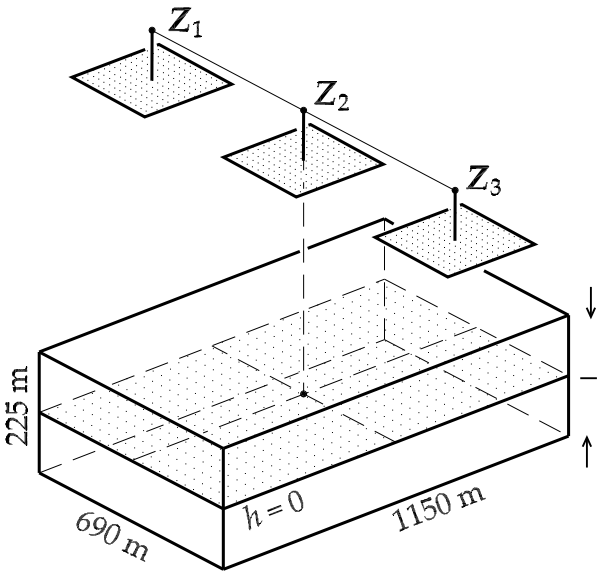

Figure 9.2: *Sketch of configuration 'Air1'. All three projection centers have* $h = 1500$*. The arrows point in direction of the cuboid compression, the limit of which is represented by the small horizontal bar.*

$$
\mathsf{P}_1 = \begin{bmatrix} 1.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & -0.000050 & 0.075000 \end{bmatrix}
$$

$$
\mathsf{P}_2 = \begin{bmatrix} 1.000000 & 0.000000 & 0.000000 & -230.000000 \\ 0.000000 & 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 0.000000 & -0.000050 & 0.075000 \end{bmatrix}
$$

$$
\mathsf{P}_3 = \begin{bmatrix} 1.000000 & 0.000000 & 0.000000 & 0.000000 \\ 0.000000 & 1.000000 & 0.000000 & -460.000000 \\ 0.000000 & 0.000000 & -0.000050 & 0.075000 \end{bmatrix}
$$

Table 9.3: *The projection matrices of configuration 'Air2'.*

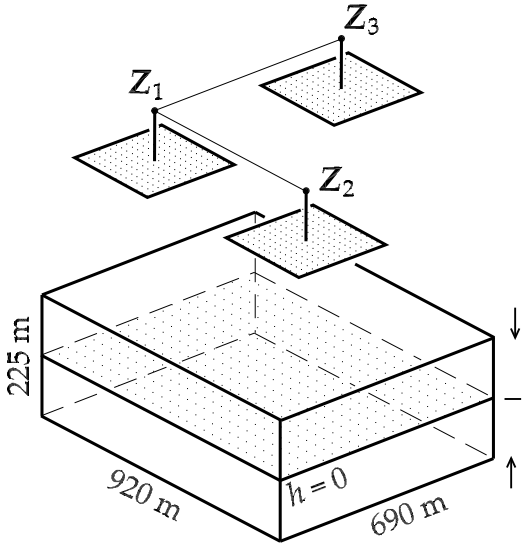

Figure 9.3: *Sketch of configuration 'Air2'. All three projection centers have* $h = 1500$*. The arrows point in direction of the cuboid compression, the limit of which is represented by the small horizontal bar.*

$$
\mathsf{P}_1 = \begin{bmatrix} 1.000000 & 0.000000 & -0.428571 & 1.071429 \\ 0.000000 & 1.000000 & 0.285714 & -6.428571 \\ 0.000000 & 0.000000 & -0.000286 & 0.003571 \end{bmatrix}
$$

$$
\mathsf{P}_2 = \begin{bmatrix} 1.000000 & 0.000000 & -0.428571 & 0.505860 \\ 0.000000 & 1.000000 & 0.285714 & -6.051526 \\ 0.000000 & 0.000000 & -0.000286 & 0.003194 \end{bmatrix}
$$

$$
\mathsf{P}_3 = \begin{bmatrix} 1.000000 & 0.000000 & -0.428571 & 0.000000 \\ 0.000000 & 1.000000 & 0.285714 & -5.714286 \\ 0.000000 & 0.000000 & -0.000286 & 0.002857 \end{bmatrix}
$$

Table 9.4: *The projection matrices of configuration 'Street1'.*

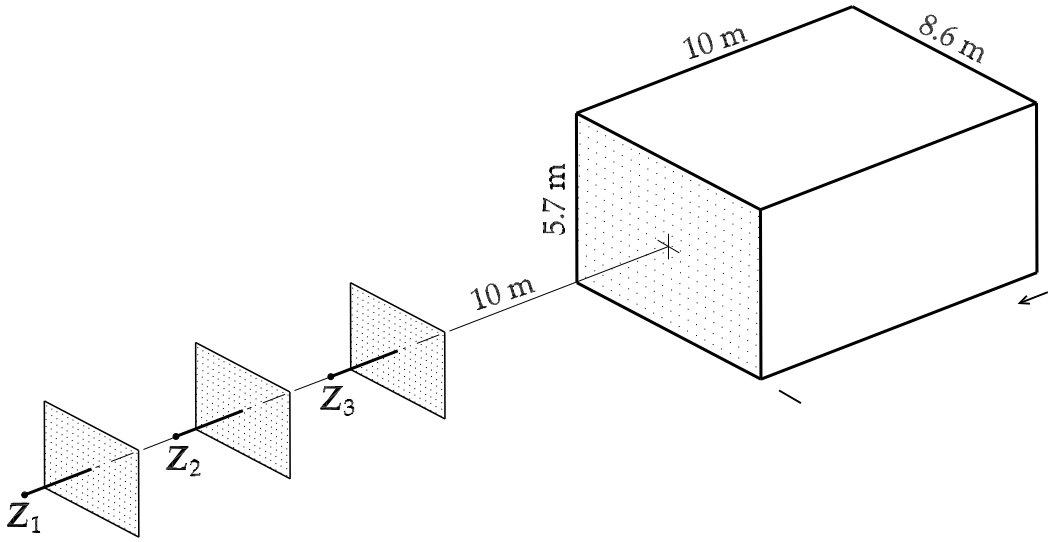

Figure 9.4: *Sketch of configuration 'Street1'. The arrows point in direction of the cuboid compression, the limit of which is represented by the small bar.*

$$
\mathsf{P}_1 = \begin{bmatrix} 1.000000 & 0.000000 & -0.428571 & 1.071429 \\ 0.000000 & 1.000000 & 0.285714 & -6.428571 \\ 0.000000 & 0.000000 & -0.000286 & 0.003571 \end{bmatrix}
$$

$$
\mathsf{P}_2 = \begin{bmatrix} 1.000000 & 0.000000 & -0.428571 & 0.505860 \\ 0.000000 & 1.000000 & 0.285714 & -6.051526 \\ 0.000000 & 0.000000 & -0.000286 & 0.003194 \end{bmatrix}
$$

$$
\mathsf{P}_3 = \begin{bmatrix} 1.000000 & 0.000000 & -0.428571 & 0.000000 \\ 0.000000 & 1.000000 & 0.285714 & -5.714286 \\ 0.000000 & 0.000000 & -0.000286 & 0.002857 \end{bmatrix}
$$

Table 9.5: *The projection matrices of configuration 'Street2'.*

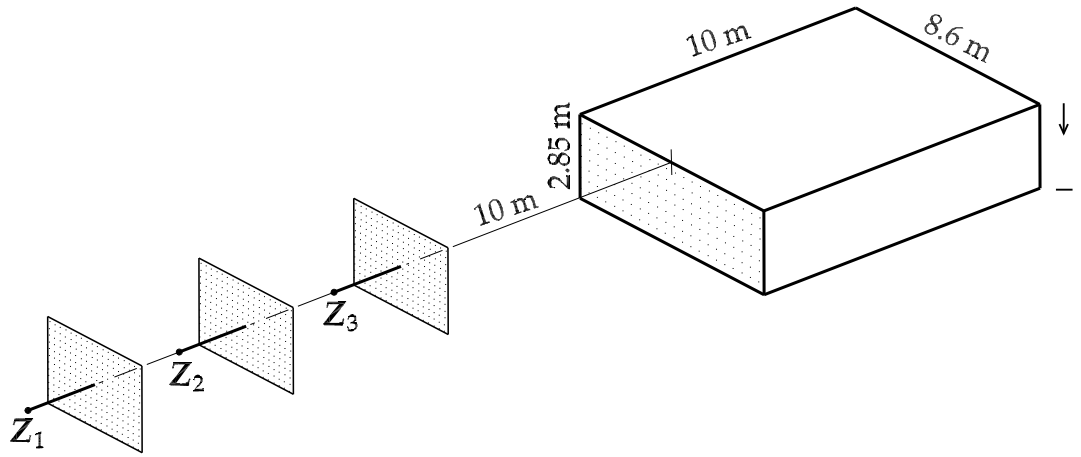

Figure 9.5: *Sketch of configuration 'Street2'. The arrows point in direction of the cuboid compression, the limit of which is represented by the small horizontal bar.*

Computation methods. For each image configuration the trifocal tensor is computed in the following ways, where minimization of reprojection error is always carried out in the Gauß-Helmert model (see section 8.3 on page 116):

'UCA': The direct linear solution or in other words the *unconstrained solution* (with 26 degrees of freedom) *minimizing algebraic error*; see section 8.1 on page 112.

'UCR': The *unconstrained solution* (with 26 degrees of freedom) *minimizing reprojection error*; see section 8.2 on page 113. This iterative estimation is initialized by the 'UCA' solution.

'CR': The *constrained solution* (with 18 degrees of freedom) *minimizing reprojection error*. The 18 degrees of freedom are realized by applying the new parameterization of section 8.3.4 on page 123 (as a *non*-minimal parameterization). This iterative estimation is initialized by the 'UCA' solution.

'CR*': This iterative estimation is identical to 'CR' but it is initialized by the known true trifocal tensor.

'CA': This is the *projection method* based on the new parameterization presented in table 8.3 on page 127. It determines that valid trifocal tensor, represented by the vector $\boldsymbol{q}$ and the new parameterization, which lies closest to the 'UCA' solution $\boldsymbol{t}$; i.e. $|\boldsymbol{t} - \boldsymbol{t}(\boldsymbol{q})| \rightarrow \min$, . The vector $\boldsymbol{q}$ found in this way also serves as initial value for the 'CR' method.

The comparison of the 'UCA' solution with the 'UCR', the 'CR' and the 'CA' solution will show the different gains in the results by minimizing reprojection error with and without considering the internal constraints as opposed to minimizing algebraic error with and without considering the constraints. This comparison will be based on the plots of the mean and maximum ground errors and on the plots of the percentage of 'bad' solutions. The first plot will show the impact of the different computations methods on the accuracies of the determined points on the ground. The second plot will show the sensitivity of the different computation methods to the compression of the cuboid.

The 'CR*' solution will act as a benchmark for assessing the critical compression situations: As it was stated above the percentage of 'bad' solutions indicates the breakdown of the tensor computation because of a too flat cuboid. Due to the random selection of the points used to determine the trifocal tensor, it may happen that those points are badly distributed in space although the cuboid itself is far away from being flat. Consequently for such badly distributed points a bad direct linear solution will result. And since the direct linear solution initializes all of the other (iterative) computation methods, chances are high that those better methods will also fail and the percentage of 'bad' solutions will be different from 0 even for non-flat cuboids.

The 'CR*' solution, however, is initialized by the true trifocal tensor of the respective image configuration, and only if this solution is also a 'bad' solution, the respective set of selected points can be regarded as being degenerate. Therefore, in the plots of the percentage of the 'bad' solutions, the respective curve of any method relying on the initialization through the direct linear solution should always be compared with the curve of the 'bad' 'CR*' solutions.

9.2 The plots of the ground errors and the percentage of failures with respect to the variation of the cuboid

9.2.1 Variation of the cuboid and 1 pixel noise in the images

In this section the plots of the averages over all 1000 samples of the mean and maximum of the ground errors, and the plots of the percentage of 'bad' solutions against the compression rate of the cuboid are depicted. The compression of the cuboids is given in percent of the camera distance. The injected noise in the images of each configuration is 1 pixel. The following thresholds for classifying 'good' and 'bad' solutions were used: 0.025 m for 'Tetra', 3.5 m for both 'Air' configurations and 1 m for both 'Street' configurations.

9.2.2 Variation of the cuboid and 10 pixel noise in the images

These plots demonstrate the effect of different noise in the images. For the configurations 'Tetra' and 'Air' 10 pixel noise is injected in the images and 10 points are used for determining the trifocal tensor. Whereas for the two 'Street' configurations 2 pixel noise is injected in the images and 25 points are used. For larger noise in the 'Street' configurations no useful results could be obtained with the chosen cuboid. The following thresholds for classifying 'good' and 'bad' solutions were used: 0.25 m for 'Tetra', 35 m for both 'Air' configurations and 1 m for both 'Street' configurations.

9.3 Investigation of the results

In the following the plots of each configuration with 1 pixel noise in the images are investigated in detail by considering the following questions:

Ground Error: What is the impact of the various computation methods, minimizing algebraic and reprojection error with or without considering the internal constraint, on the mean and maximum ground errors?

Flatness: At which *minimum thickness* of the object points do the various computation methods still return a reasonable estimate of the trifocal tensor, before the determination fails?

A summary of these investigations and general formulas are given in the section 9.4 on page 163.

We expect the following from the plots: For a small number of point correspondences the ground errors should be larger and the minimum thickness should be reached earlier, than for a large number of point correspondences. To show this behavior we will start with the smallest possible point number, i.e. 7, and increase this number step by step. However, only the configurations 'Tetra' and 'Air2', which have a strong geometry, return acceptable results already with 7 points, whereas the smallest possible number for 'Air1' was 8, and for the two 'Street' configurations, due to their very weak geometry, even for 10 points very bad results were obtained.

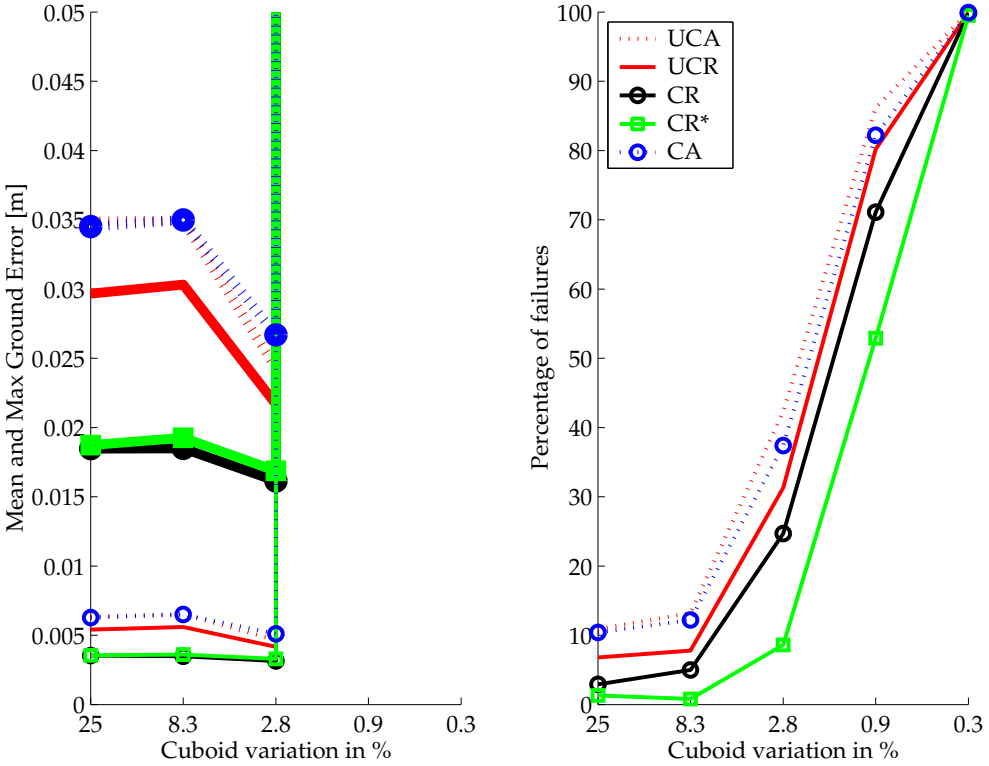

Figure 9.6: *Configuration: 'Tetra', 7 points, 1 pixel noise, camera distance: 3 m, threshold for bad solutions: 0.025 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

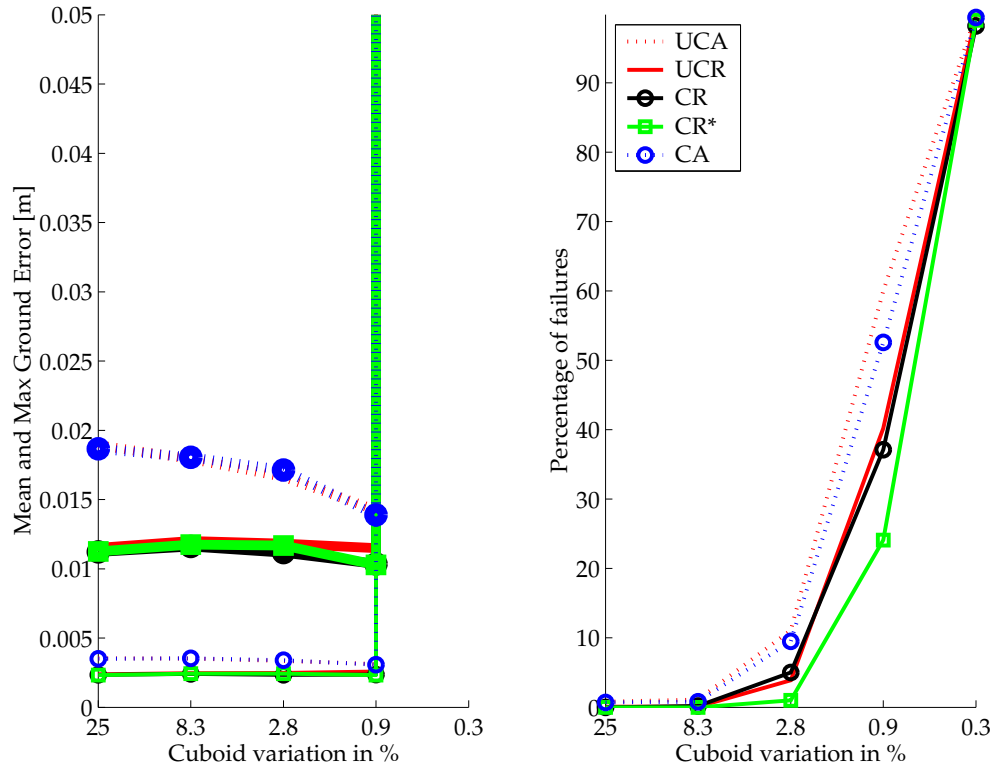

Figure 9.7: *Configuration: 'Tetra', 8 points, 1 pixel noise, camera distance: 3 m, threshold for bad solutions: 0.025 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

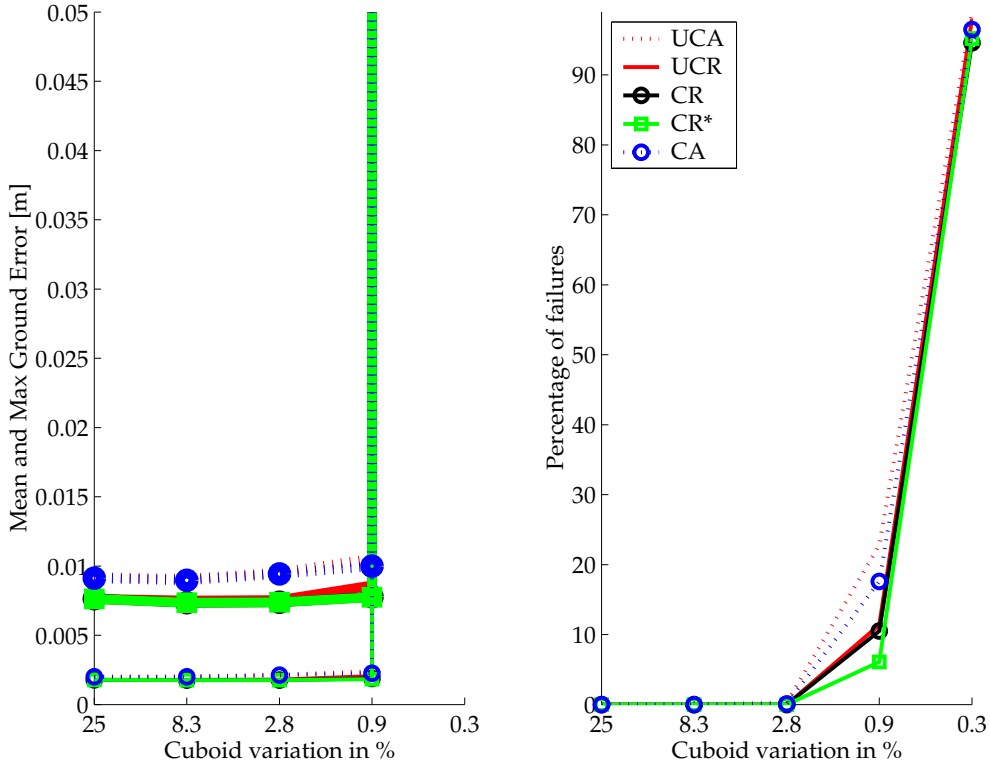

Figure 9.8: *Configuration: 'Tetra', 10 points, 1 pixel noise, camera distance: 3 m, threshold for bad solutions: 0.025 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

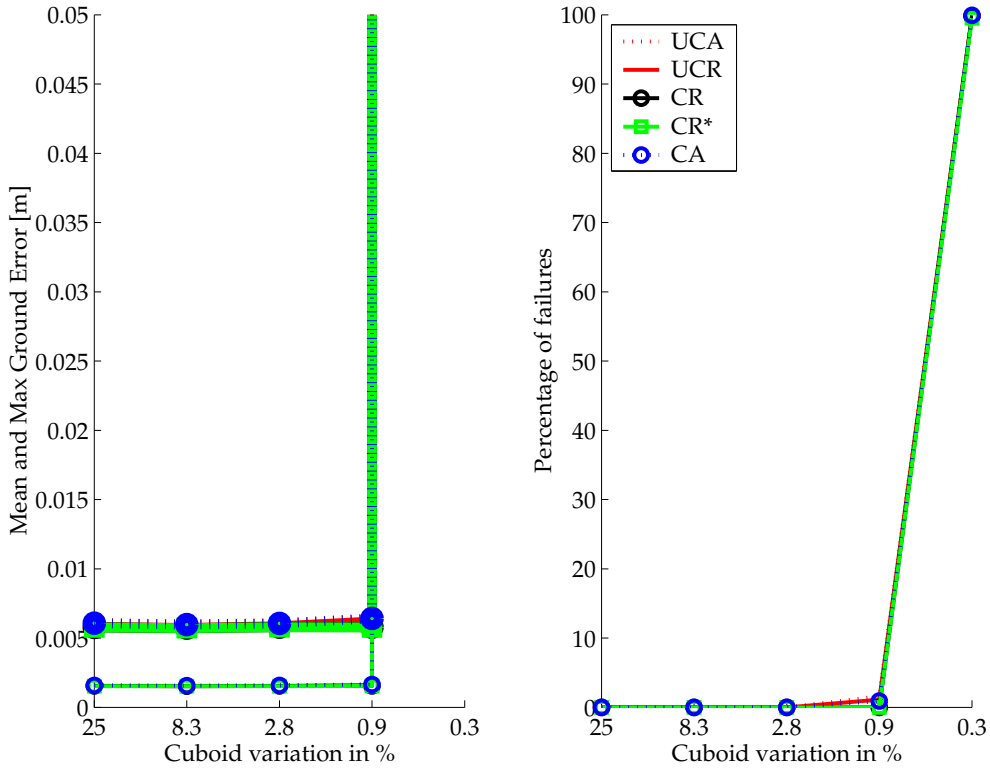

Figure 9.9: *Configuration: 'Tetra', 15 points, 1 pixel noise, camera distance: 3 m, threshold for bad solutions: 0.025 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

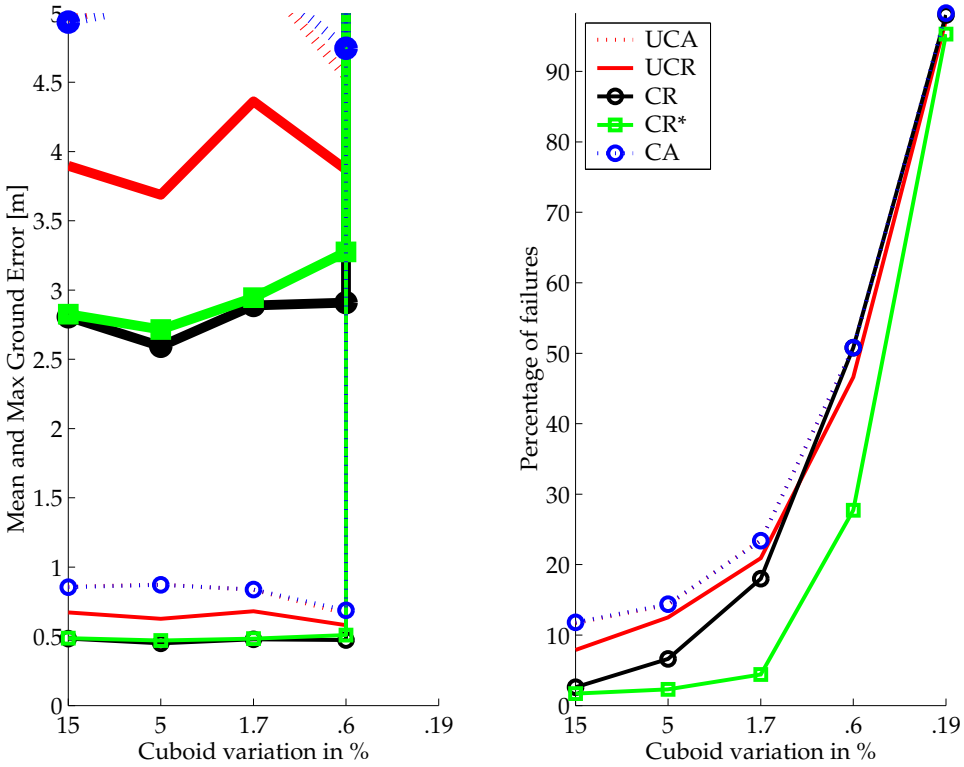

Figure 9.10: *Configuration: 'Air1', 8 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

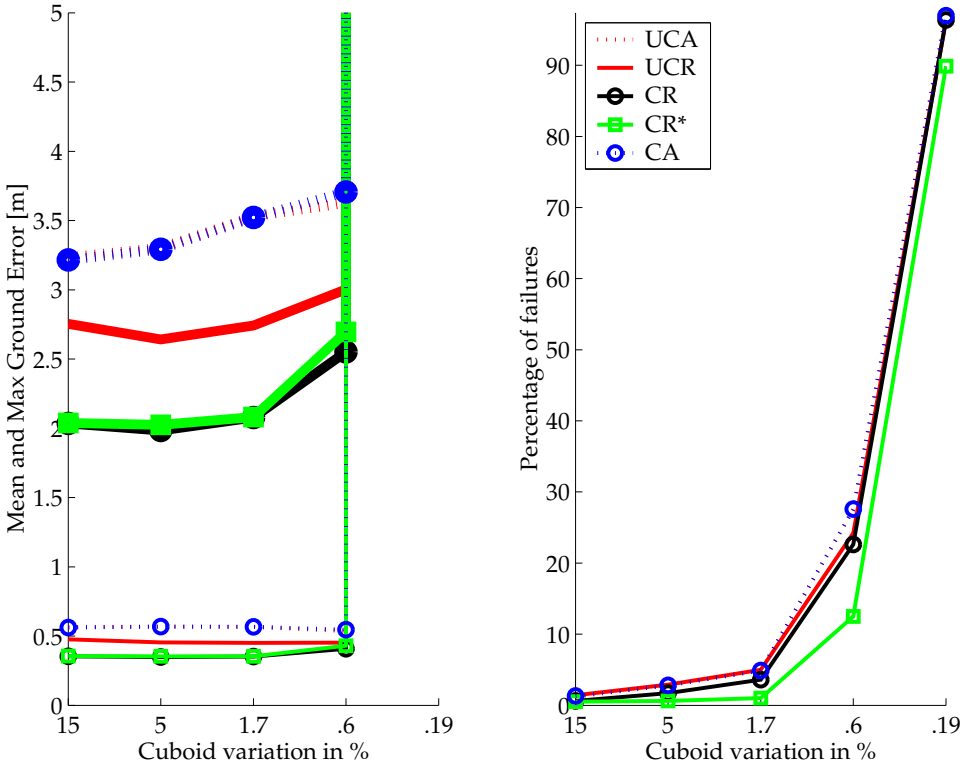

Figure 9.11: *Configuration: 'Air1', 9 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

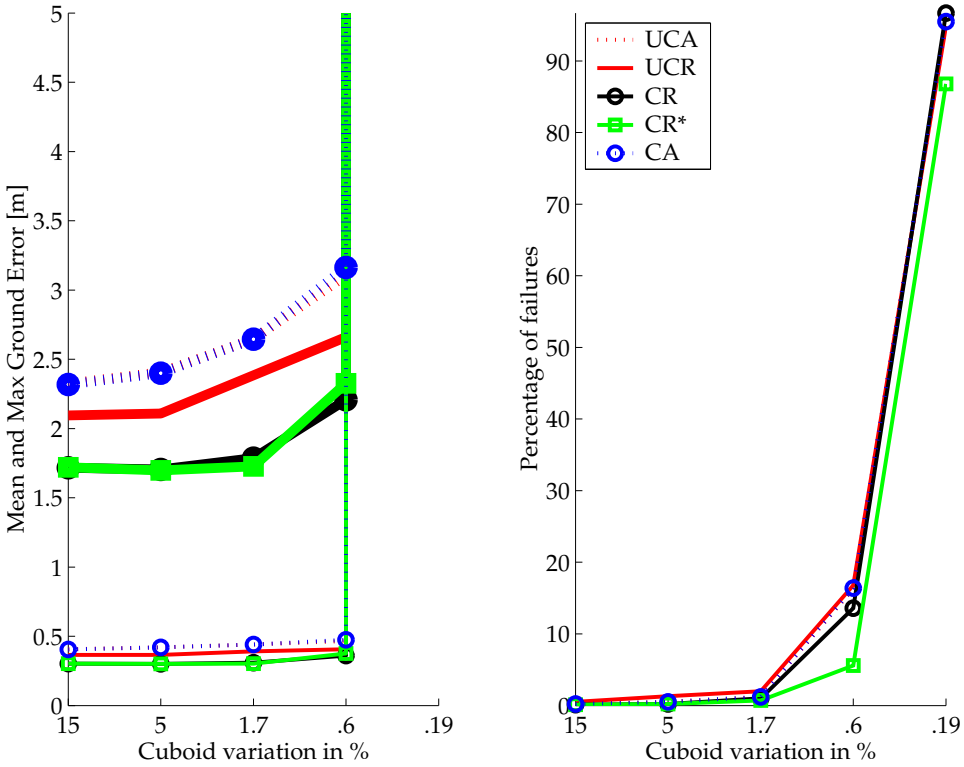

Figure 9.12: *Configuration: 'Air1', 10 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

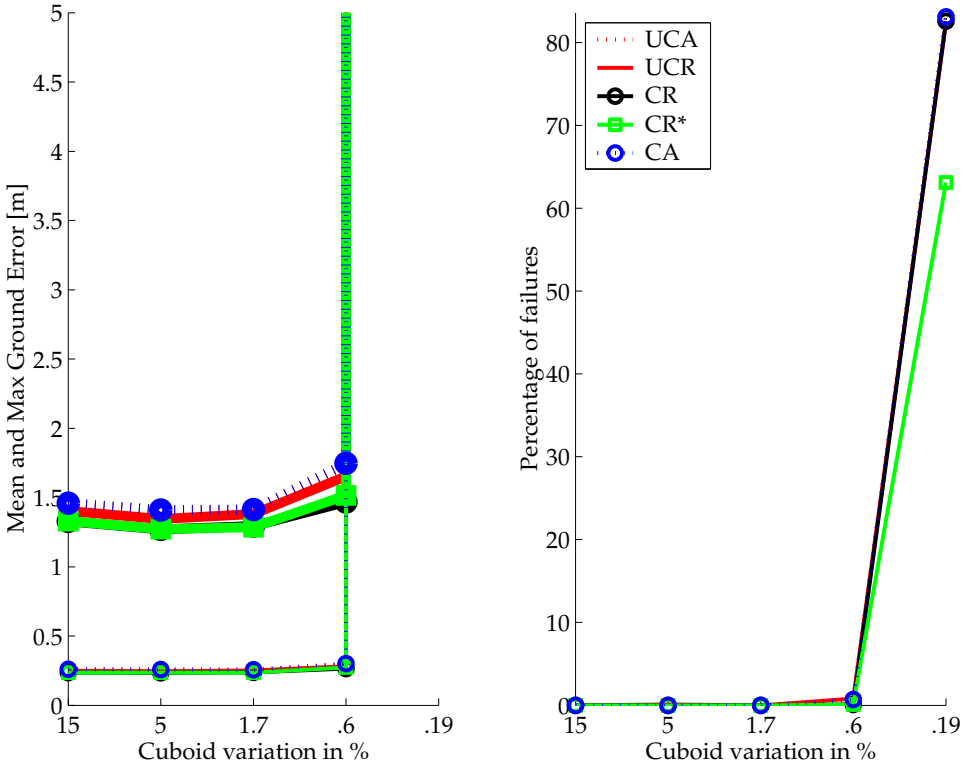

Figure 9.13: *Configuration: 'Air1', 15 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

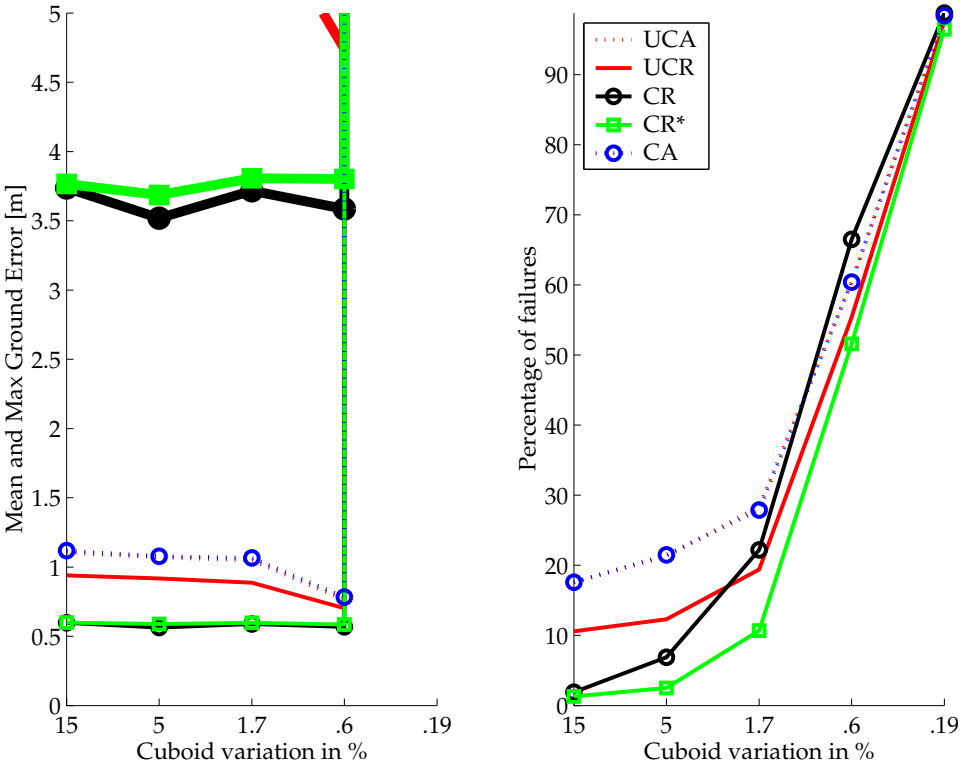

Figure 9.14: *Configuration: 'Air2', 7 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

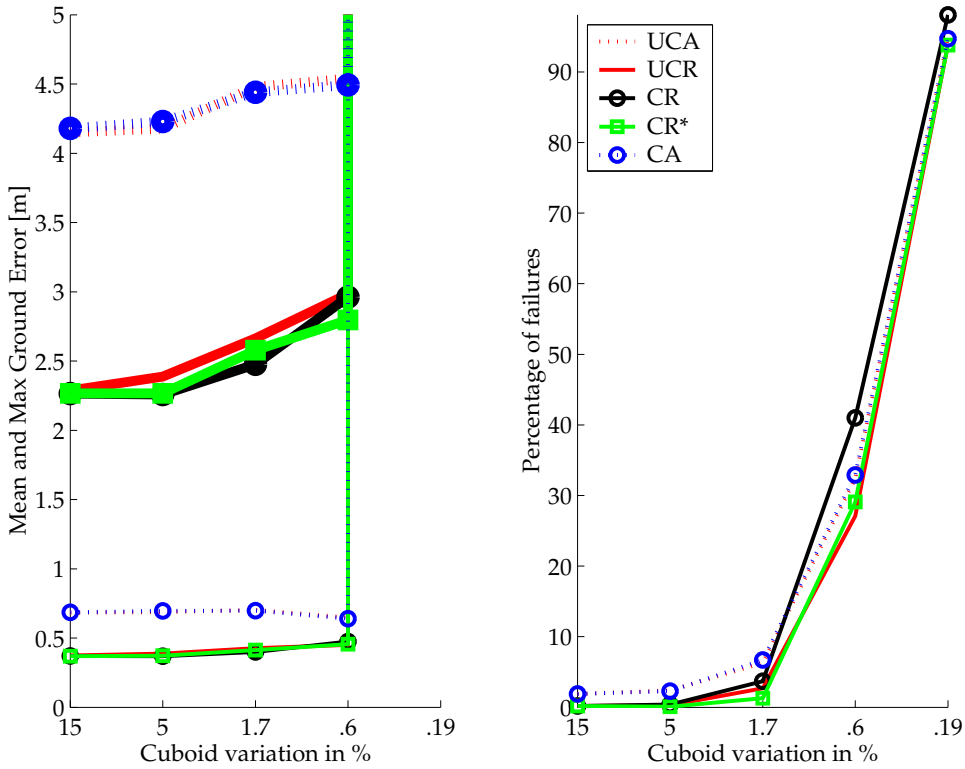

Figure 9.15: *Configuration: 'Air2', 8 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

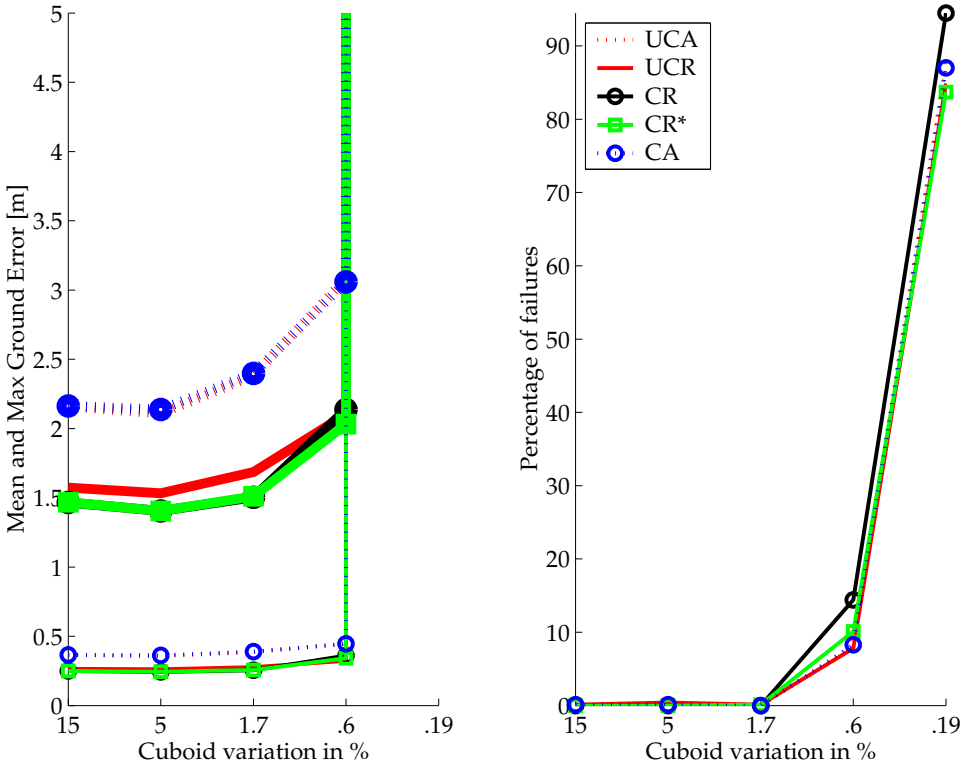

Figure 9.16: *Configuration: 'Air2', 10 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

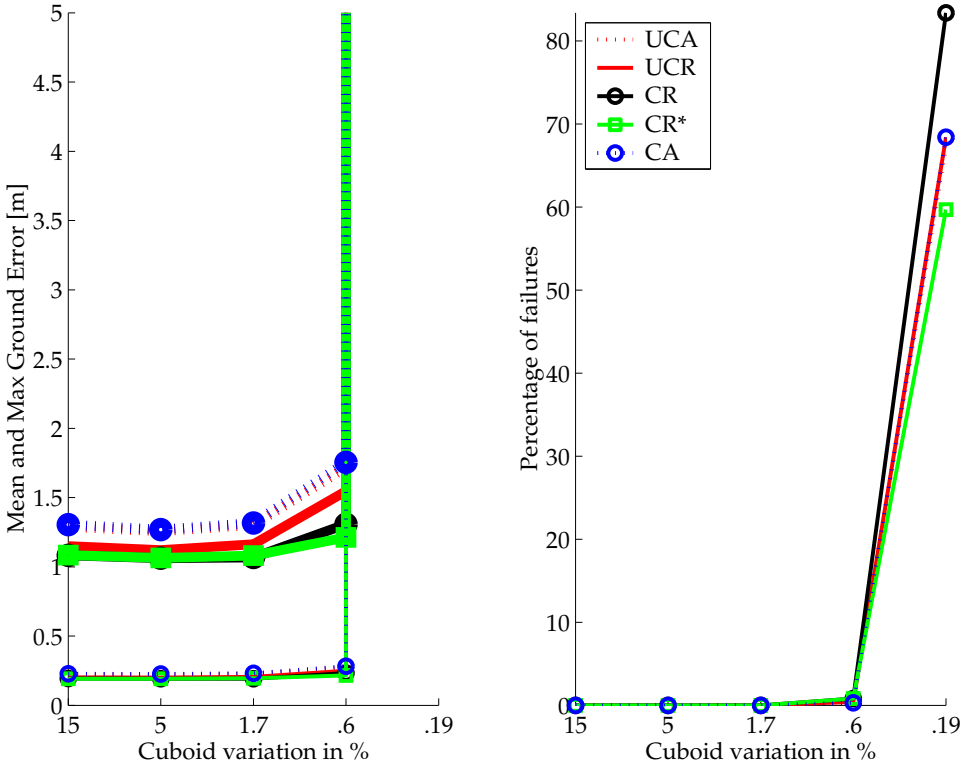

Figure 9.17: *Configuration: 'Air2', 15 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

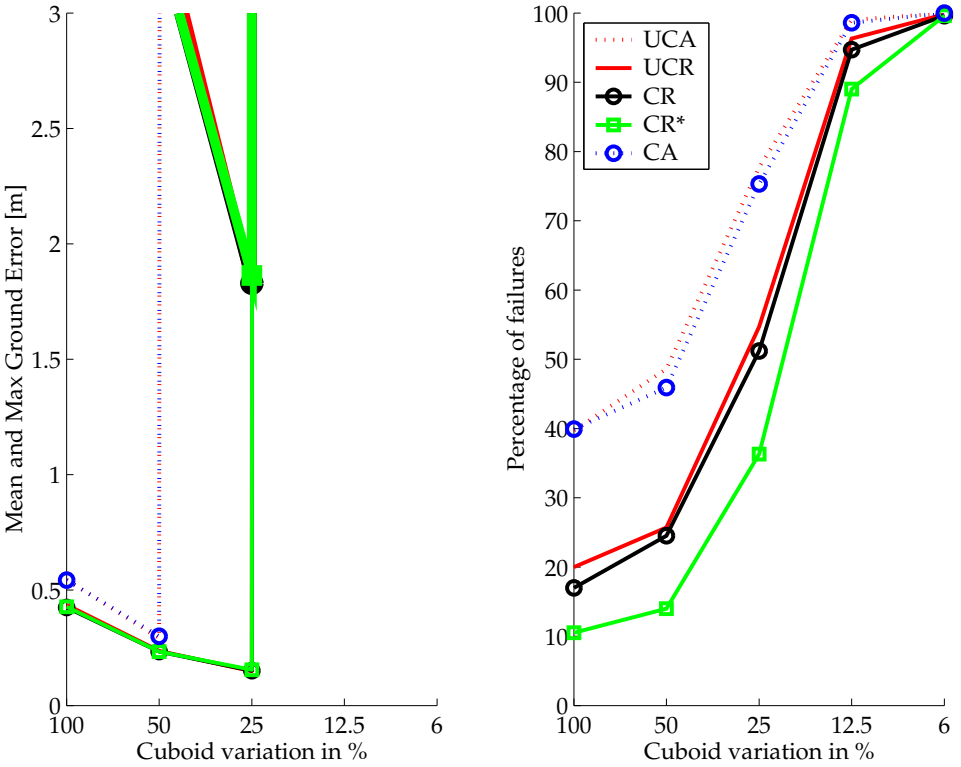

Figure 9.18: *Configuration: 'Street1,' 10 points, 1 pixel noise, camera distance: 10 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

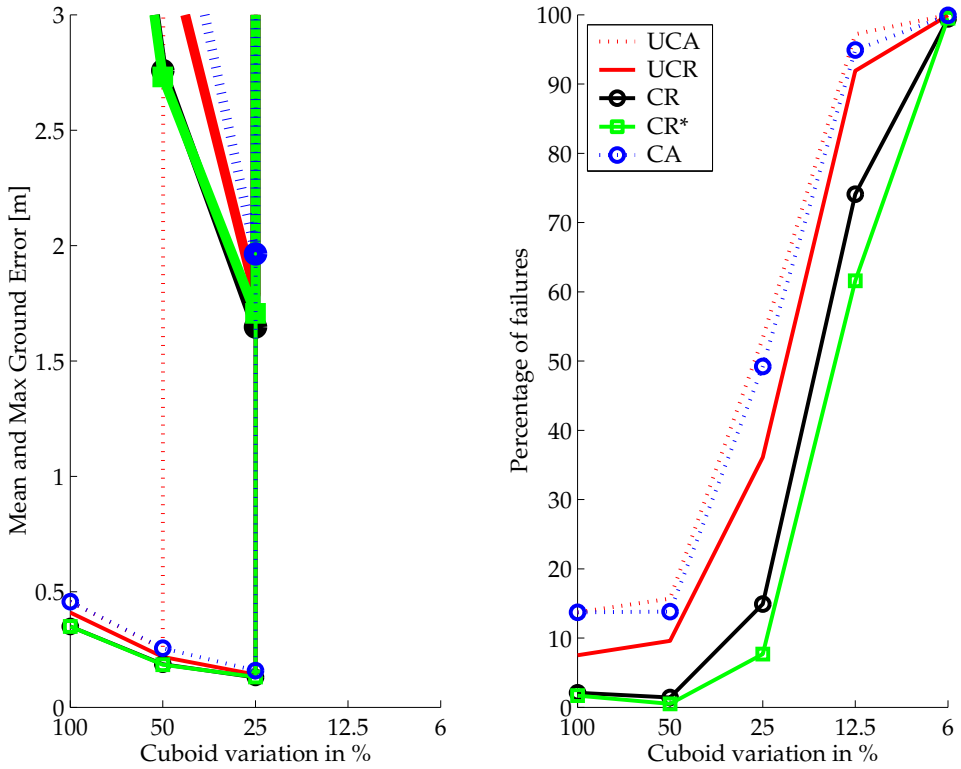

Figure 9.19: *Configuration: 'Street1,' 15 points, 1 pixel noise, camera distance: 10 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

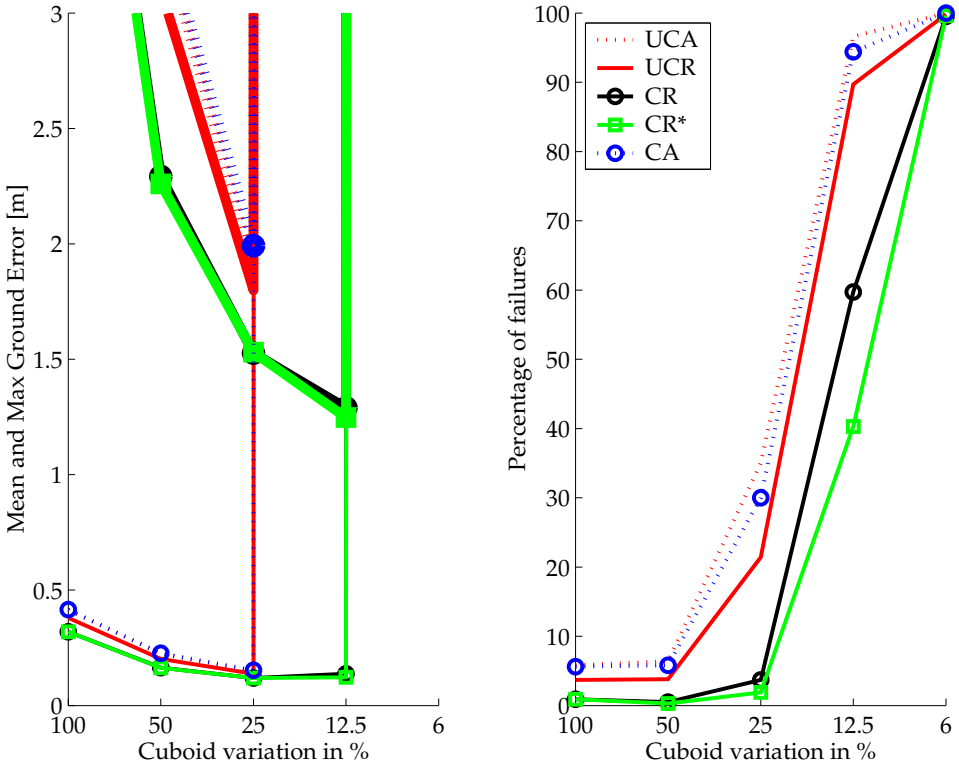

Figure 9.20: *Configuration: 'Street1,' 20 points, 1 pixel noise, camera distance: 10 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

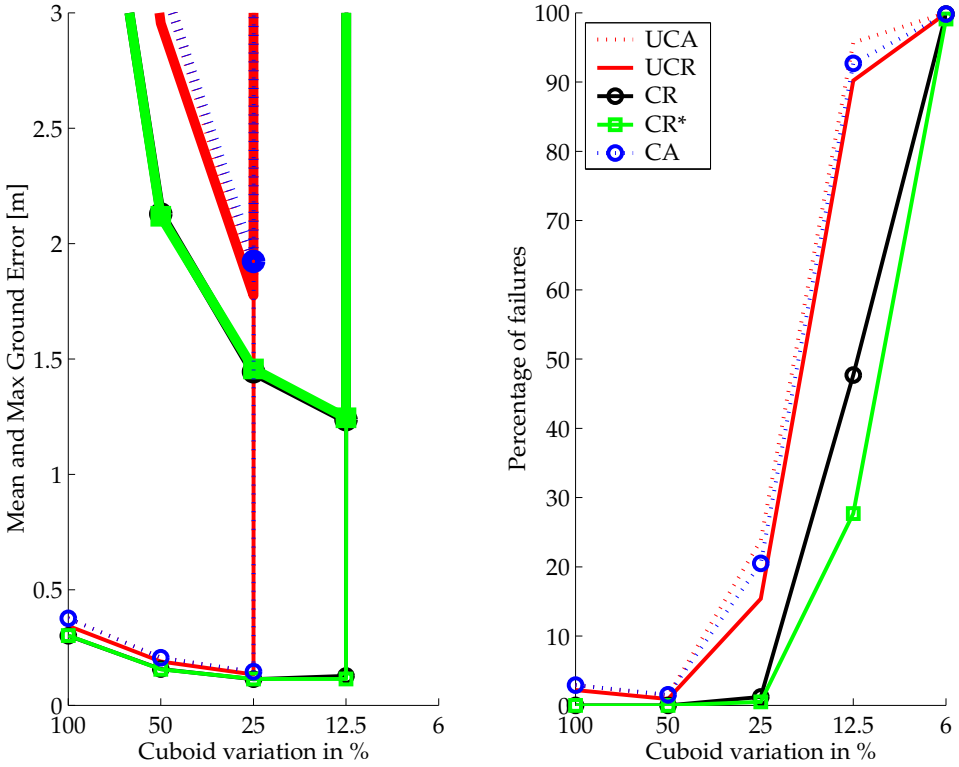

Figure 9.21: *Configuration: 'Street1,' 25 points, 1 pixel noise, camera distance: 10 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

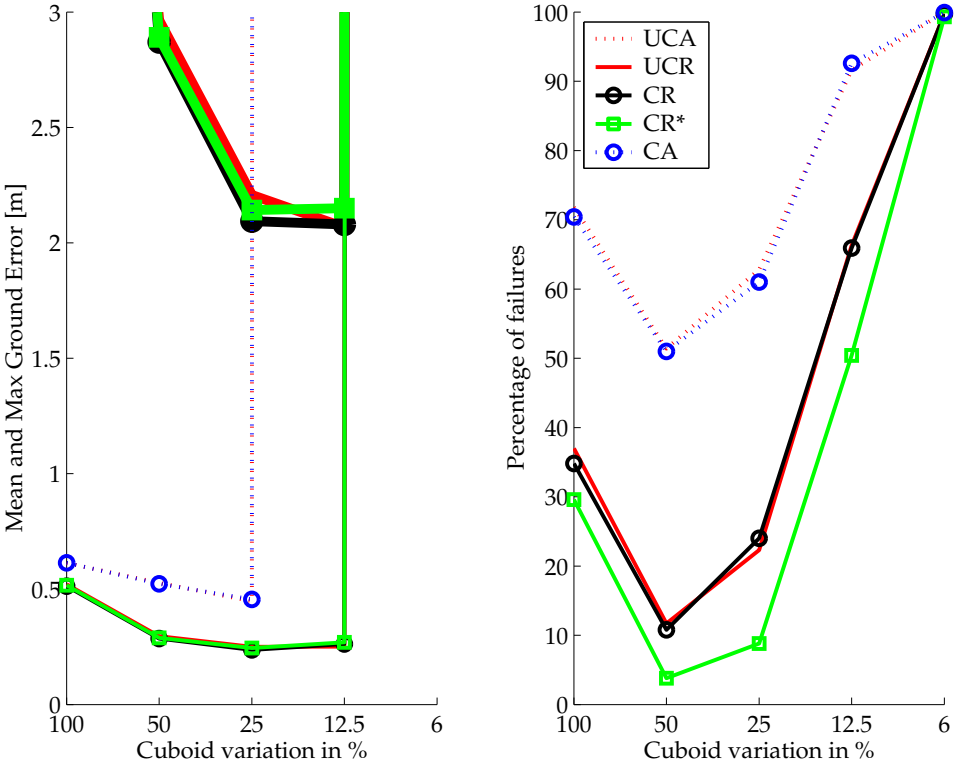

Figure 9.22: *Configuration: 'Street2,' 10 points, 1 pixel noise, reference distance: 5.7 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

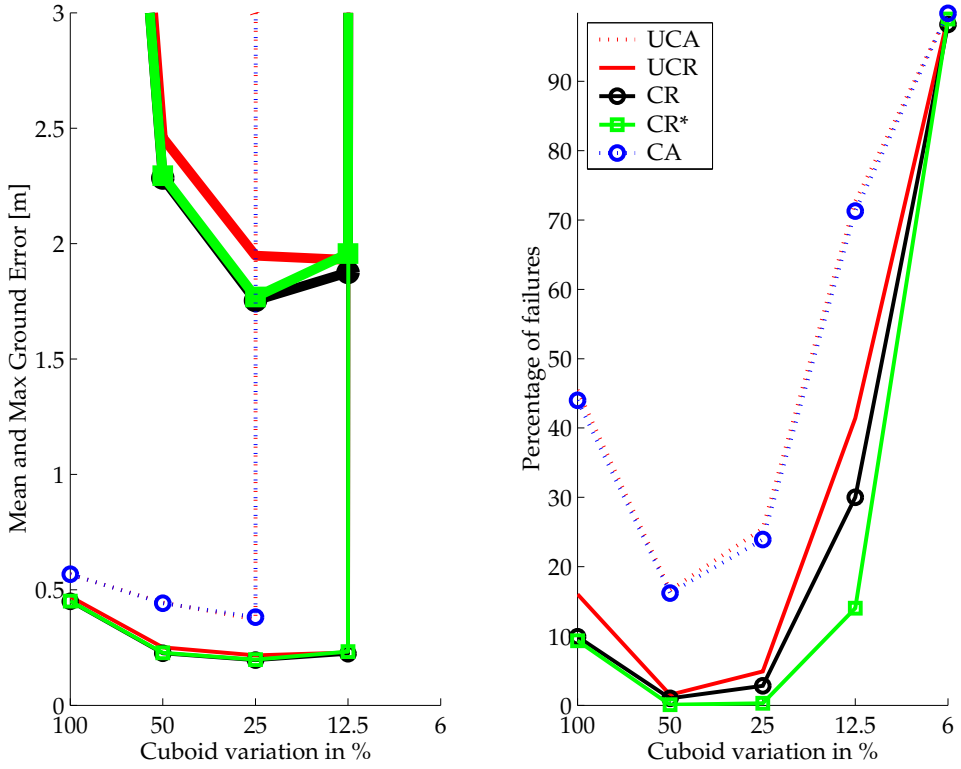

Figure 9.23: *Configuration: 'Street2,' 15 points, 1 pixel noise, reference distance: 5.7 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

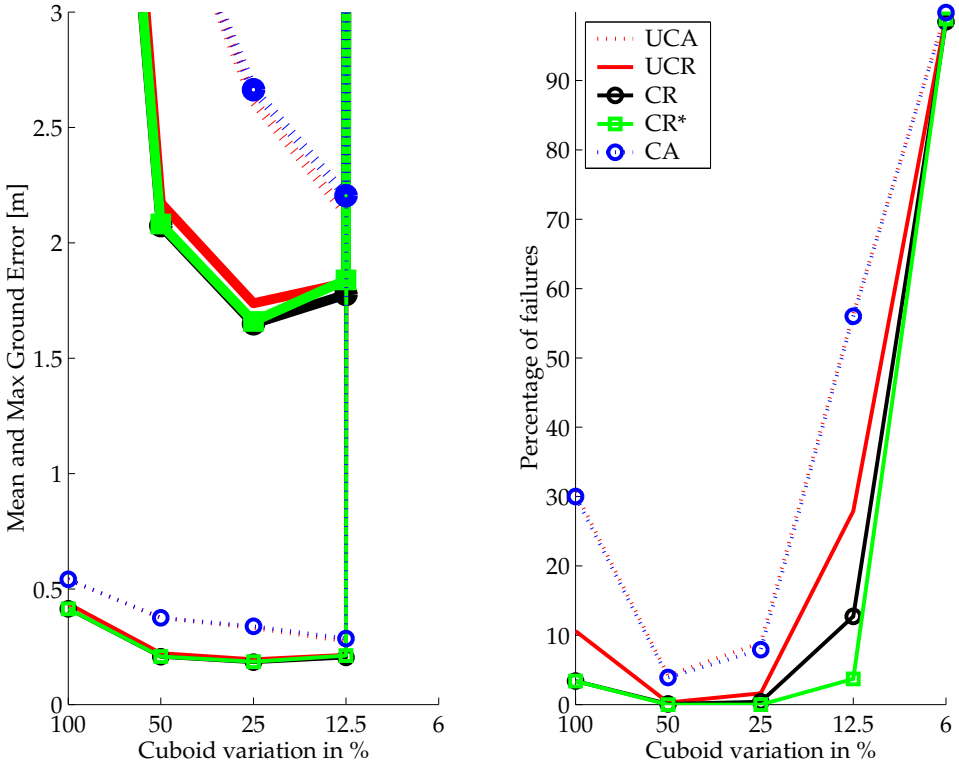

Figure 9.24: *Configuration: 'Street2,' 20 points, 1 pixel noise, reference distance: 5.7 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

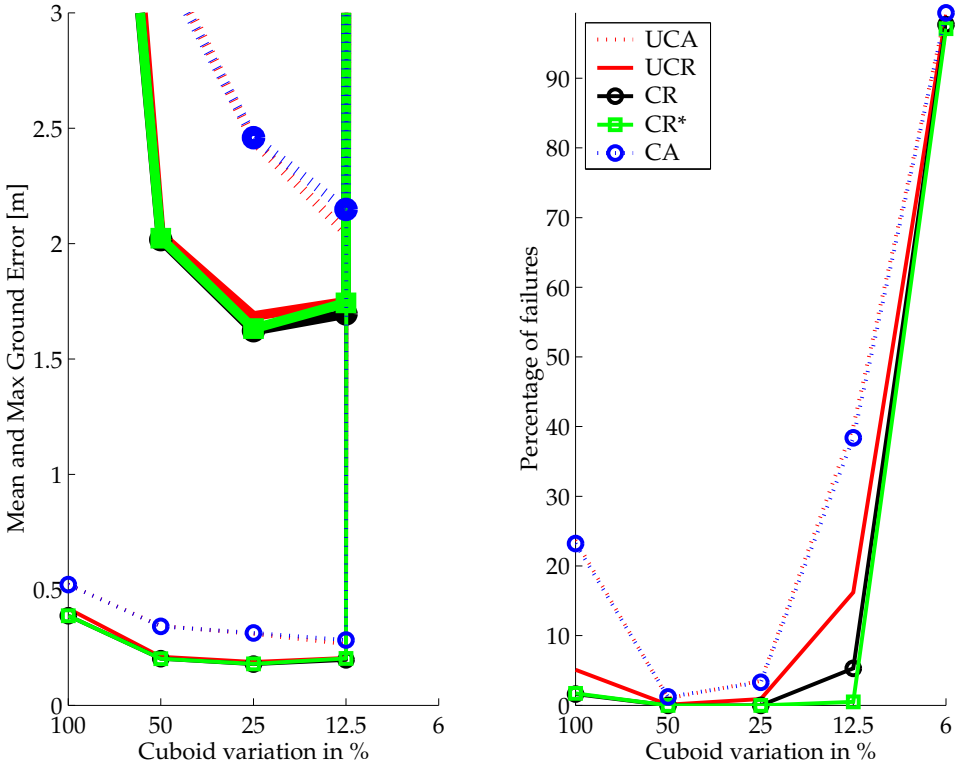

Figure 9.25: *Configuration: 'Street2,' 25 points, 1 pixel noise, reference distance: 5.7 m, threshold for bad solutions: 1 m. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

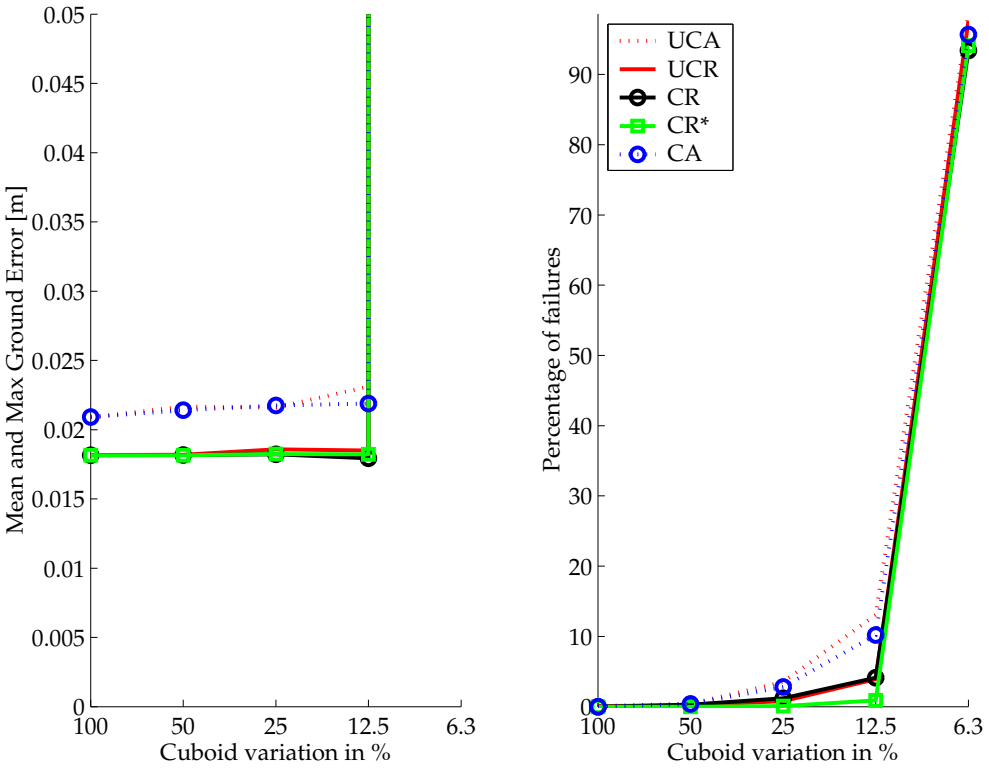

Figure 9.26: *Configuration: 'Tetra', 10 points, 10 pixel noise, camera distance: 3 m, threshold for bad solutions: 0.25 m.; cf. figure 9.8. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

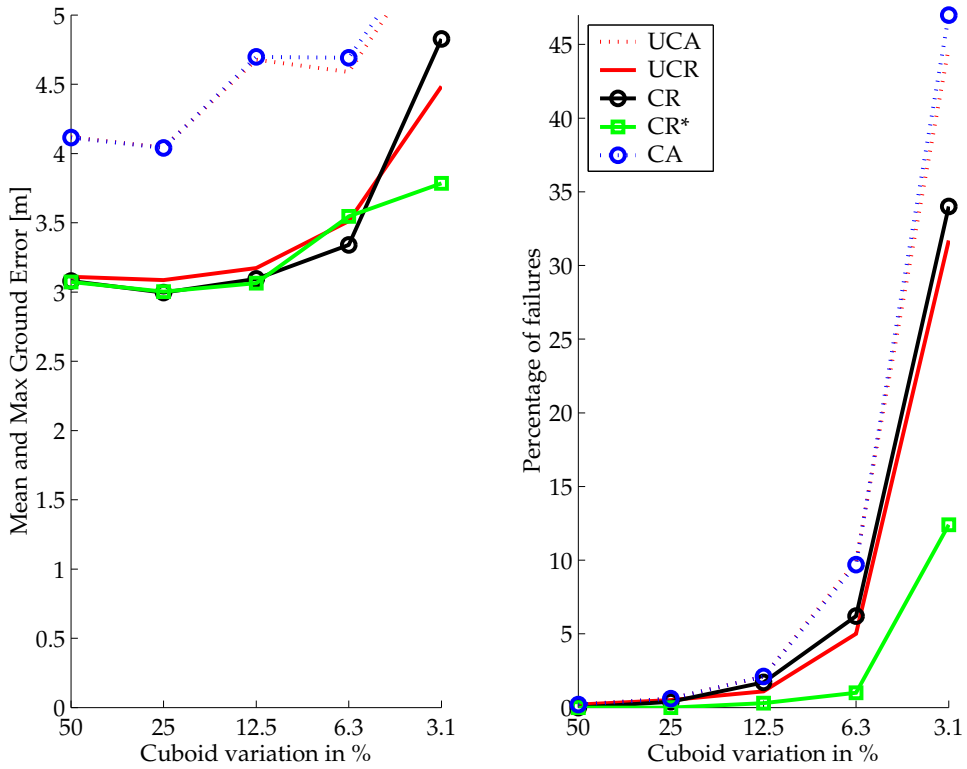

Figure 9.27: *Configuration: 'Air1', 10 points, 10 pixel noise, camera distance: 1500 m, threshold for bad solutions: 35 m; cf. figure 9.12. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

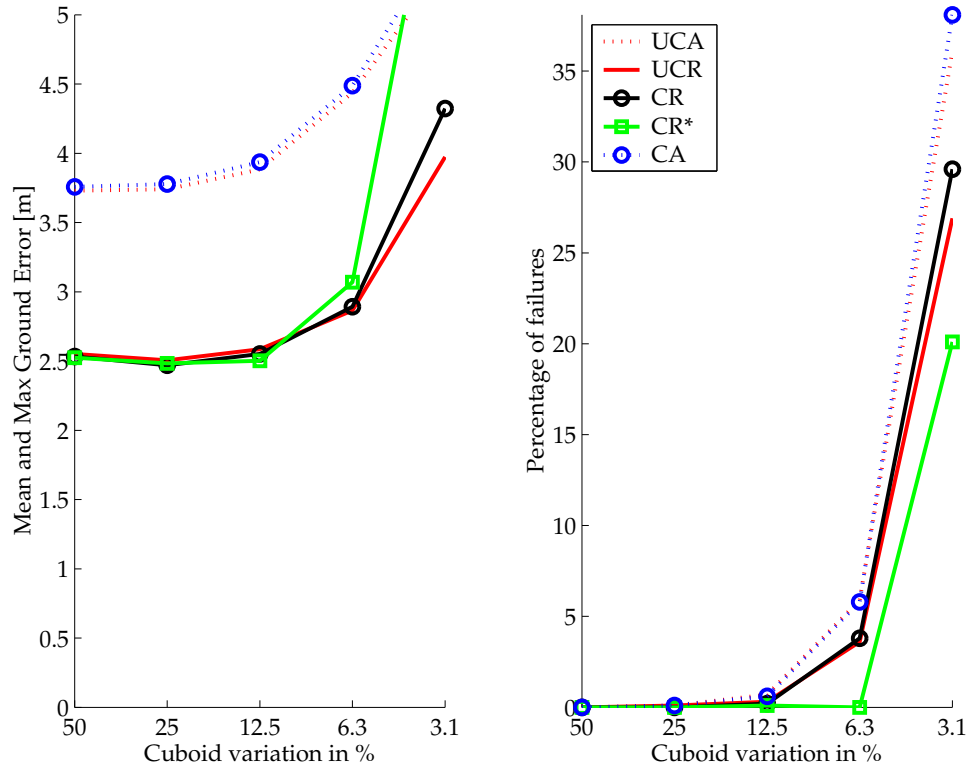

Figure 9.28: *Configuration: 'Air2', 10 points, 10 pixel noise, camera distance: 1500 m, threshold for bad solutions: 35 m; cf. figure 9.16. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

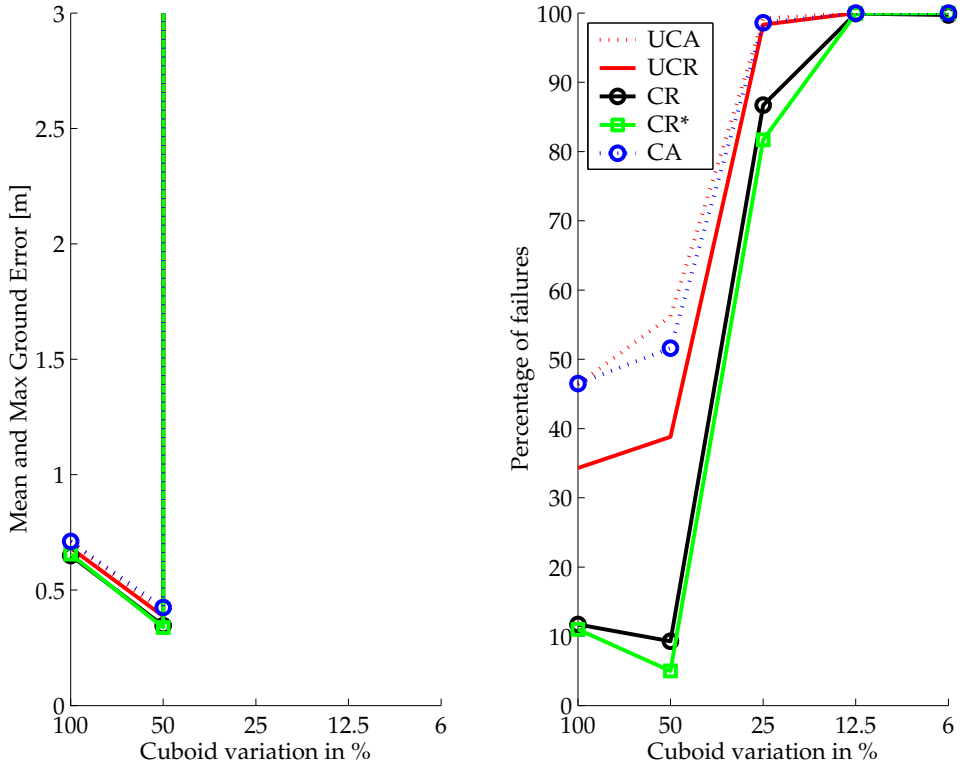

Figure 9.29: *Configuration: 'Street1,' 25 points, 2 pixel noise, camera distance: 10 m, threshold for bad solutions: 1 m; cf. figure 9.21. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

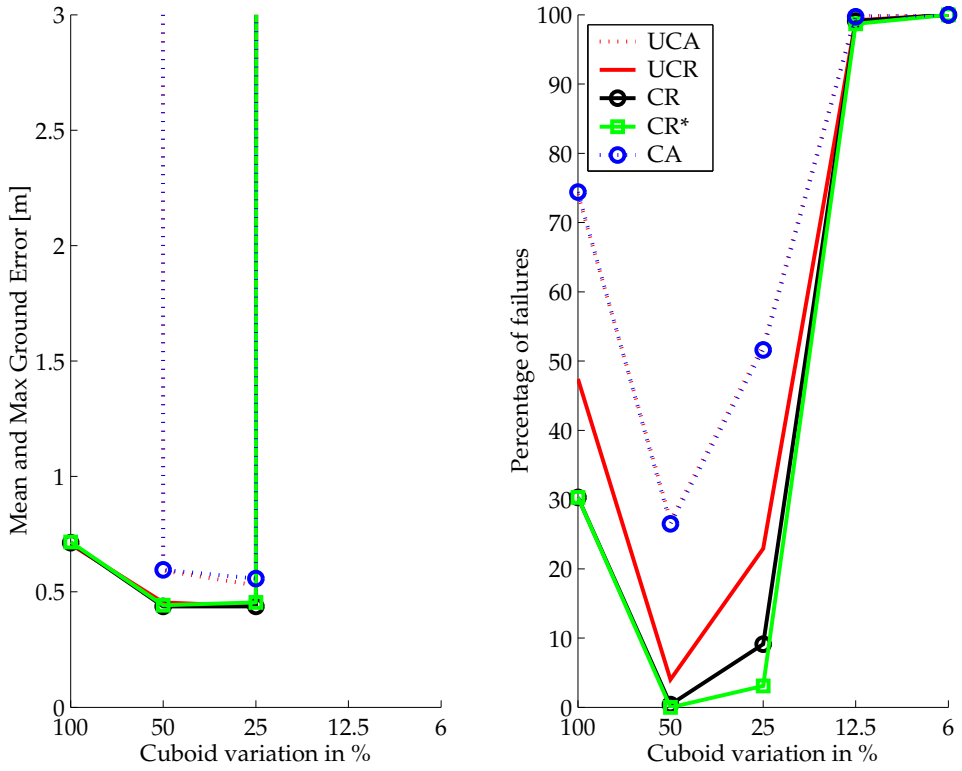

Figure 9.30: *Configuration: 'Street2,' 25 points, 2 pixel noise, reference distance: 2.85 m, threshold for bad solutions: 1 m; cf. figure 9.25. The mean ground error is plotted with thin lines, the maximum ground error with thick lines.*

The algebraic minimizations, 'UCA' and 'CA', should in general produce larger ground errors and should breakdown earlier than the reprojection methods, 'UCR' and 'CR'. Methods that neglect the internal constraints of the tensor should also perform worse than the methods that consider them. Therefore the direct linear solution 'UCA' is expected to be at the low end, and the constrained method minimizing reprojection error 'CR' is expected to be at the high end of the performances.

Further, we compare the plots with 1 pixel noise in the images of section 9.2.1 with the corresponding plots with larger noise of section 9.2.2.

9.3.1 Configuration 'Tetra'

The figures 9.6 – 9.9 on the pages 145 and 146 show the plots of the configuration 'Tetra' based on image noise of 1 pixel. The cuboid compressions are given in percent of the camera distance of 3 m.

Ground Error: Looking at the left plot in each of these figures, we see that up including 8 points the unconstrained computation minimizing reprojection error 'UCR' performs better than any algebraic computation ('UCA' and 'CA'); which indicates the benefit of minimizing reprojection error. From 10 points onwards all five computation methods ('UCA', 'UCR', 'CR', 'CR*', 'CA') produce practically the same curves for the mean and maximum ground errors; i.e. even the simple direct linear solution 'UCA' performs equally well to the constrained method 'CR' minimizing reprojection error.

Flatness: Looking at the right plot in each of these figures, we see that from 7 points onwards the percentage of bad solutions is slightly smaller for the reprojection methods 'UCR' and 'CR' than for the algebraic methods 'UCA' and 'CA', whereas there is almost no difference in the bad solutions between considering or neglecting the constraints. For 7 points and flatness 25% the reprojection methods initialized by the direct linear solution deviates from the 'CR*' solution, which is initialized by the true tensor; this indicates that the direct linear solution can be insufficient to initialize the iterative methods. For 10 points and 15 points all methods show the same minimum thickness of 2.8% and 0.9% respectively.

Noise in the images: By comparing the plots based on 10 points with 1 and 10 pixel noise in the images (figure 9.8 on page 146 and figure 9.26 on page 155), we see that the mean and maximum ground errors and the minimum thickness of the computations are practically scaled by the ratio of the noise in the images. Therefore, at least for the noise-range considered here, the results depend linearly on the noise in the images.

9.3.2 Configuration 'Air1'

The figures 9.10 – 9.13 on the pages 9.10 on page 147 and 9.13 on page 148 show the plots of the configuration 'Air1' based on image noise of 1 pixel. The cuboid compressions are given in percent of the camera distance of 1500 m.

Ground Error: Looking at the left plot in each of these figures, we see that up including 9 points the unconstrained computation minimizing reprojection error 'UCR' performs better than any algebraic computation ('UCA' and 'CA'); which indicates the benefit of minimizing reprojection error. From 10 points onwards all five computation methods ('UCA', 'UCR', 'CR', 'CR*', 'CA') produce practically the same curves for the mean ground errors; i.e. even the simple direct linear solution 'UCA' performs equally well to the constrained method 'CR' minimizing reprojection error. The maximum ground error for the algebraic methods is still larger, but from 15 points onwards mean and maximum error for all methods are identical.

Flatness: Looking at the right plot in each of these figures, we see that for 8 points the curves of the bad solutions for all methods have values different from zero even for 15% flatness. For 8 and 9 points all methods relying on the initialization through the direct linear solution 'UCA' have larger percentage of bad solutions than the 'CR*' method. Again, this shows that for small number of points the 'UCA' can be insufficient to initialize the iterative methods. From 10 points onwards, however, all methods are identical in their behavior concerning the bad solutions and show the same minimum thickness, which is 5%, 1.7% and 0.6% for 9, 10 and 15 points respectively.

Noise in the images: By comparing the plots based on 10 points with 1 and 10 pixel noise in the images (figure 9.12 on page 148 and figure 9.27 on page 156), we see that the mean and maximum ground errors and the minimum thickness of the computations are practically scaled by the ratio of the noise in the images. Therefore, at least for the noise-range considered here, the results depend linearly on the noise in the images.

9.3.3 Configuration 'Air2'

The figures 9.14 – 9.17 on the pages 9.14 on page 149 and 9.17 on page 150 show the plots of the configuration 'Air2' based on image noise of 1 pixel. The cuboid compressions are given in percent of the camera distance of 1500 m.

Ground Error: Looking at the left plot in each of these figures, we see that up including 10 points the unconstrained computation minimizing reprojection error 'UCR' performs better than any algebraic computation ('UCA' and 'CA'); which indicates again the benefit of minimizing reprojection error. From 10 points onwards all five computation methods ('UCA', 'UCR', 'CR', 'CR*', 'CA') produce practically the same curves for the mean ground errors; i.e. even the simple direct linear solution 'UCA' performs equally well to the constrained method 'CR' minimizing reprojection error. The maximum ground error for the algebraic methods is still larger, but from 15 points onwards mean and maximum error for all methods are identical. So, except that 'Air2' performs better than 'Air1' for small number of points, from 10 points onwards both 'Air' configurations perform identically.

Flatness: Looking at the right plot in each of these figures, we see that for 7 points the curves of the bad solutions are to some extent the same for all methods, except for 'CR*', which is significantly better; this indicates again that the direct linear solution for seven points can be

not sufficient to initialize the iterative methods. From 10 points onwards, however, all methods are identical in their behavior concerning the bad solutions and show the same minimum thickness, which then only depends on the number of point correspondences.

Noise in the images: By comparing the plots based on 10 points with 1 and 10 pixel noise in the images (figure 9.16 on page 150 and figure 9.28 on page 156), we see that the mean and maximum ground errors and the minimum thickness of the computations are practically scaled by the ratio of the noise in the images. Therefore, at least for the noise-range considered here, the results depend linearly on the noise in the images.

9.3.4 Configuration 'Street1'

The figures 9.18 – 9.21 on the pages 9.18 on page 151 and 9.21 on page 152 show the plots of the configuration 'Street1' based on image noise of 1 pixel. The cuboid compressions are given in percent of the camera distance of 10 m.

Ground Error: Looking at the left plot in each of these figures, we see that from 10 points onwards all methods show practically the same curves of mean and maximum error; so there is practically no gain in accuracy in computing 'CR' compared to 'UCA'.

Flatness: The right plot of the bad solutions, however, shows a radically different picture as the percentage of bad solutions for the two algebraic methods 'UCA' and 'CA' always lie clearly above all reprojection methods; although the difference gets smaller with increasing number of points.

Noise in the images: By comparing the plots based on 25 points with 1 and 2 pixel noise in the images (figure 9.21 on page 152 and figure 9.29 on page 157), we see that the mean and maximum ground errors of the computations are to some extent scaled by the ratio of the noise in the images. This also holds true for the percentage of bad solutions. However, for 1 pixel noise the smallest percentage of bad solutions is 0% (for a cuboid compression of 25%), whereas for 2 pixel noise the smallest percentage is about 10% (for a cuboid compression of 50%). This shows that due to the weak image geometry the impact of the image noise on the critical planarity is much stronger than for the previous configurations.

Additional note: The reason for almost identical ground error curves between algebraic and reprojection methods is caused by the large number of 10 – 25 points, and not due to some hidden properties of the algebraic methods. For smaller number of points the algebraic methods are practically useless concerning accuracy. Since we can see that even for 10 points any method produces bad solutions with a percentage of at least 20 %, the high number of points was selected to show their requirement for a reliable, i.e. not bad, solution. So the ground error plots of the algebraic methods should be understood as: *When the algebraic solution is good, it is comparable to the reprojection method, however, the chances of getting a good result from an algebraic method is much smaller than for a reprojection method*.

Note also the strange course of the ground error plots, which gets better with increasing flatness. This behavior is explained by the way the cuboid is compressed. This is done in *direction to the images* by fixing the side of the cuboid, which faces the images. Therefore all points move closer to the images and the *mean intersection angle of corresponding projection rays gets larger*. Consequently the accuracy of the determined points also becomes better.

9.3.5 Configuration 'Street2'

The figures 9.22 – 9.25 on the pages 9.22 on page 153 and 9.25 on page 154 show the plots of the configuration 'Street2' based on image noise of 1 pixel. The cuboid compressions are given in percent of a reference distance of 2.85 m; see figure 9.5.

Ground Error: Looking at the left plot in each of these figures, we see that from 10 points onwards all reprojection methods are clearly distinct from the algebraic methods (see also the following additional note 1), while there is no difference between the constrained and unconstrained methods of the same minimization.

Flatness: The plot of the bad solutions shows, that – similar to 'Street1' – the percentage of bad solutions for the two algebraic methods 'UCA' and 'CA' always lie clearly above all reprojection methods; although the difference gets smaller with increasing number of points.

Noise in the images: By comparing the plots based on 25 points with 1 and 2 pixel noise in the images (figure 9.25 on page 154 and figure 9.30 on page 157), we see that, similar to 'Street1', the mean and maximum ground errors of the computations are also to some extent scaled by the ratio of the noise in the images. This also holds true for the percentage of bad solutions. Opposed to 'Street1', however, the smallest percentage of bad solutions is 0% for both 1 and 2 pixel noise; namely for the cuboid compression of 25% and 50% respectively. For 2 pixel noise on both sides of this minimal thickness of 50% the percentage of bad solutions quickly raises. So here again due to the weak geometry the impact of the image noise on the critical planarity is stronger than for the configurations with stable geometry; i.e. 'Tera' and 'Air'.

Additional note 1: From the way the ground error curve of the algebraic methods progresses from 10 to 25 points without ever reaching the curve of the reprojection methods, we see that there is a clear bias between these two kinds of minimizations. This is opposed to 'Street1', where already for 10 points practically no difference between both minimizations could be observed.

This bias is caused by the nature of the algebraic error, which is the deviation from zero on the right side of the homogenous point-trilinearities. For each corresponding triple of image points $\{\mathbf{x}', \mathbf{x}'', \mathbf{x}'''\}$ *four* point-trilinearities (7.29) on page 79 are set up. Each of these four trilinearities can be represented as $\mathbf{x}'^{\top}\boldsymbol{\lambda}'_{jk} = 0,\ j,k = 1,2$. The line $\boldsymbol{\lambda}'_{jk}$ in image Ψ_1 is the projection of a 3D line, which is created by the intersection of the back-projection of the lines $\boldsymbol{\sigma}''_j$ and $\boldsymbol{\sigma}'''_k$; cf. section 7.3 and figure 7.7 on page 79.

Due to the special geometry of the 'Street'-configurations, the lines $\boldsymbol{\lambda}'_{12}$ and $\boldsymbol{\lambda}'_{21}$ run approximately radially from the principal point of image Ψ_1 to the point $\mathbf{x}'$. This holds true for image

Ψ_1 in both configurations 'Street1' and 'Street2'. The bias in the algebraic minimization for 'Street1' is caused by the distribution of the points $\mathbf{x}'$ in image Ψ_1, which is totaly different to 'Street2'.

In 'Street1' the cuboid is compressed in direction of the basis of the images, therefore the distribution area of the points $\mathbf{x}'$ stays practically the same for all cuboid compressions. In 'Street2' the cuboid is compressed orthogonally to the basis of the images, in direction to the street, therefore the distribution area of the points $\mathbf{x}'$ is limited to the lower half of image Ψ_1; see figure 9.31.

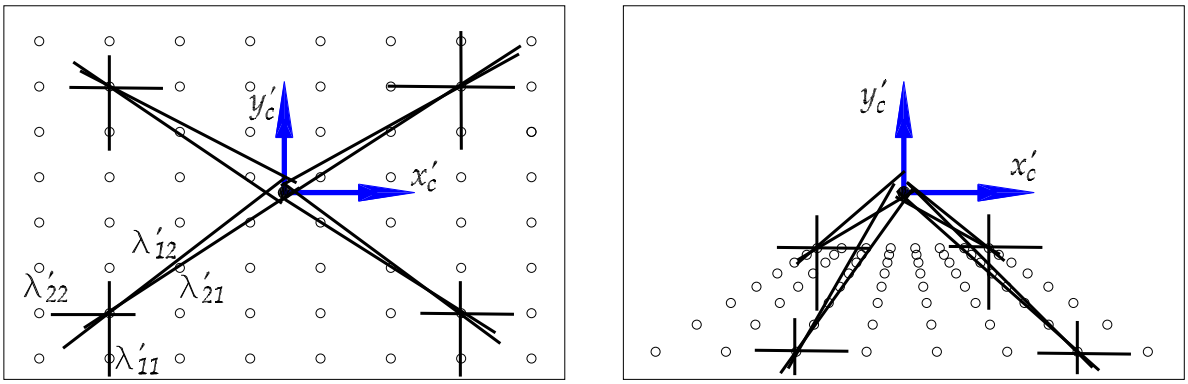

Figure 9.31: *The image Ψ_1 of configuration 'Street1' (left) and 'Street2' (right), with the lines $\boldsymbol{\lambda}'_{jk}$ drawn in four representative points. In image Ψ_1 of 'Street1' the lines $\boldsymbol{\lambda}'_{12}$ and $\boldsymbol{\lambda}'_{21}$ are not limited in their direction, whereas in image Ψ_1 of 'Street2' these lines, and partly the lines $\boldsymbol{\lambda}'_{11}$, have all approximately the same direction (i.e. the y'_c-direction).*

During the algebraic minimization the deviations on the right side of the trilinearities $\mathbf{x}'^{\top}\boldsymbol{\lambda}'_{jk} = 0$ are minimized. These trilinearities represent the incidence relation between the point $\mathbf{x}'$ and the line $\boldsymbol{\lambda}'_{jk}$. Therefore the right side of these trilinearities corresponds to the orthogonal distance of the point $\mathbf{x}'$ to the line $\boldsymbol{\lambda}'_{jk}$, however, scaled by the norm of the homogenous part of $\boldsymbol{\lambda}'_{jk}$. Now, as we can see from figure 9.31 for the configuration 'Street2', the lines $\boldsymbol{\lambda}'_{12}$, $\boldsymbol{\lambda}'_{21}$, and partly the lines $\boldsymbol{\lambda}'_{11}$, have all approximately the same direction (i.e. the y'_c-direction). Therefore the contribution of the error component orthogonal to this direction (i.e. the x'_c-direction) to the total algebraic error is almost three times higher than for the y'_c-direction.

These different, unjustified, weights of the four point-trilinearities of 'Street2' are the reason for this bias. For 'Street1' the lines $\boldsymbol{\lambda}'_{12}$ and $\boldsymbol{\lambda}'_{21}$ are not limited in their direction as the points $\mathbf{x}'$ are scattered in the whole image area.

The algebraic error corresponds to the distance of the point $\mathbf{x}'$ to the line $\boldsymbol{\lambda}'_{jk}$, scaled by the norm of the homogenous part of $\boldsymbol{\lambda}'_{jk}$. The scale of $\boldsymbol{\lambda}'_{jk}$ depends on the coordinates of the points $\mathbf{x}''$ and $\mathbf{x}'''$ in the other images. So there could be an additional source of bias in the algebraic error due to the scale difference between the Euclidian coordinates, which can be some 1000 pixels, and the homogenous extension of 1. Because of the conditioning following section 3.7 on page 28, however, this bias should be eliminated to a large extent, as all image points are translated to their centroid and are appropriately scaled afterwards. This conjecture is also backed by the results of 'Street1', which does not show this kind of bias.

Additional note 2: A strange course of the ground error plots, which gets better with increasing flatness, can be observed, similar to 'Street1'. The reason for this course, however, is different to 'Street1', where it was caused by the increasing mean intersection angle of corresponding projection rays due to the cuboid compression. Here the mean intersection angle stays more or less the same, since the depth of the cuboid remains fixed and only its height is flattened. For 'Street2' the improvement is caused by the increasing distance of the points to the common basis of the three images, which cuts right through the middle of the uncompressed cuboid. For object points on this basis the intersection angle of corresponding projection rays is zero and therefore the object point is undeterminable[1]. According to this, in the vicinity to the basis the intersection angle will be very small, resulting in very badly determined object points. The percentage of such points decreases as the cuboid for 'Street2' is compressed. Consequently the mean and maximum ground error get smaller.

Since the mean intersection angle for 'Street2' does not increase that much as it does for 'Street1', the actual values of the ground errors is higher than for 'Street1'.

9.4 Summary

The expectations stated in the beginning of section 9.3 on page 144 are practically confirmed, however the direct linear solution performed much better than presumed. We can summarize the results of the examples with 1 pixel noise in the images in the following way:

- The *ground errors* obtained for the different computation methods tend to be the same
  - for a particular image configuration, if the *number of point correspondences increases*, and
  - for a particular number of point correspondence, if the *stability of the image geometry increases*.

  Therefore, from 10 points onwards for 'Tetra', 'Air1' and 'Air2', and from 25 points onwards for 'Street1' and 'Street2', all computation methods return practically the same result; i.e. then *the direct linear solution is equivalent to the constrained computation minimizing reprojection error*.

- For small numbers of point correspondences, ≤ 8 for the stable configurations 'Tetra', 'Air1' and 'Air2' and for the unstable configurations 'Street1' and 'Street2' in general, the unconstrained minimization of reprojection error generally performed better than any algebraic (constrained or unconstrained) minimization. Therefore, *the benefit of minimizing only reprojection error, is larger than of considering only the constraints*.

- The algebraic minimizations can produce a bias for certain lopsided configurations, i.e. 'Street2', due to the incorrect weighting of the trilinearities.

- Concerning the impact of the *minimum thickness of the cuboid* on the various computation methods, we can say

[1] For the preparation of these examples we were aware of this fact and therefore the number of points place in the cuboid along each direction is *even*, because for an odd number, points exactly on the basis would have been created.

  - for image configurations with strong geometry, i.e. 'Tetra', 'Air1' and 'Air2', the minimum thickness is practically independent on the computation method of the tensor, and becomes smaller for larger numbers of point correspondences.
  - for image configurations with weak geometry, i.e. 'Street1' and 'Street2', the minimum thickness is larger for the algebraic methods and smaller for the reprojection methods, where the consideration of the internal constraints adds a significant additional benefit; and for larger numbers of point correspondences the minimum thickness also gets smaller generally.

- Depending on the number of point correspondences, the computation of the trifocal tensor was *still successful* for the minimum thickness of the cuboid given in table 9.6; afterwards it failed. Because of the *discrete* cuboid compressions the actual minimum thickness is smaller than the presented values; i.e. if the percentage of 'bad' solutions for one compression is zero and for the next it is non-zero, then the actual minimum thickness lies somewhere in between.

configuration	number of points		
	8	10	15
'Tetra' camera distance approx. 3 m	8.3 %	2.8 %	0.9 %
'Air1' camera distance approx. 1500 m	>15 %	1.7 %	0.6 %
'Air2' camera distance approx. 1500 m	5 %	1.7 %	0.6 %
configuration	number of points		
	15	20	25
'Street1' camera distance approx. 10 m	50 %	50 %	25 %
'Street2' reference distance approx. 2.85 m	50 %	25 %	12.5 %

Table 9.6: *Minimum thickness in percent of the camera distance for which the computation of the trifocal tensor was still successful. For the configurations 'Tetra', 'Air1' and 'Air2' the given values hold for any computation method, whereas for the two 'Street' configurations they hold only for the constrained method minimizing reprojection error.*

By comparing the plots for the configurations 'Tetra', 'Air1' and 'Air2' based on 1 pixel noise in the images with the respective plots based on 10 pixel noise, we see that the ground errors and the minimum thickness of the computations are practically scaled by the ratio of the injected noise. Therefore, using 10 as a representative number of points for the configurations 'Tetra', 'Air1' and 'Air2', we can summarize the mean and maximum ground error and the minimum thickness with the following formulas (distance D to object, noise σ in the image, principal distance c):

'Tetra':

$$\begin{aligned} \text{Mean ground error} &= 2\,\frac{D\cdot\sigma}{c} \\ \text{Max ground error} &= 9\,\frac{D\cdot\sigma}{c} \\ \text{Minimum thickness} &\leq 98\,\frac{D\cdot\sigma}{c} \end{aligned}$$

'Air1':

$$\begin{aligned} \text{Mean ground error} &= 4\,\frac{D\cdot\sigma}{c} \\ \text{Max ground error} &= 23\,\frac{D\cdot\sigma}{c} \\ \text{Minimum thickness} &\leq 340\,\frac{D\cdot\sigma}{c} \end{aligned}$$

'Air2':

$$\begin{aligned} \text{Mean ground error} &= 3\,\frac{D\cdot\sigma}{c} \\ \text{Max ground error} &= 20\,\frac{D\cdot\sigma}{c} \\ \text{Minimum thickness} &\leq 340\,\frac{D\cdot\sigma}{c} \end{aligned}$$

We can use the symbol '$\leq$', because of the *discrete* cuboid compressions. Also note that the principal distance is only included in these formulas to be independent on the units in the image space, and that the formulas may not be very correct for totally different principal distances for the same cameras. Here the principal distance for 'Air1' and 'Air2' is 20000 pixel and for 'Tetra' it is 3500 pixel.

Similar formulas for the two 'Street' configuration are not given, since due to the very weak geometry the ground error depends highly on the cuboid compression, with better ground accuracies for thinner cuboids. Further, since for higher noise levels than 2 pixel no useful results were obtained, it can be deduced that the influence of the noise level is more crucial, which further prevents the derivation of general formulas.

Since the 'Air1' configuration is probably the most relevant for practical applications, we can take a closer look at the distribution of the planer and height components of the errors of this configuration; using 10 points and 1 pixel noise. From figure 9.32 on the following page we see that the mean planar error and the mean height error are about (12 cm/50 cm) for the 'CR' method and (17 cm/68–75 cm) for the 'UCA' method. For the same configuration but using *calibrated* instead of uncalibrated images, we can adapt the formulas of [Kraus 1993] and obtain the planar error as $1[\text{pixel}]\cdot 1500[\text{m}]/20000[\text{pixel}] \mathrel{\widehat{=}} 7.5$ cm and the height error as $0.15‰\cdot 1500[\text{m}] \mathrel{\widehat{=}} 23.5$ cm. Therefore, the constrained determination of the trifocal tensor is approximately 2 times worse and the direct linear solution is approximately 3 times worse. In general, however, the relative orientation of the three images obtained from both methods, especially from the direct linear solution, is a sufficient good approximation to initialize a final bundle block adjustment during which the known interior orientation of the images can be fully exploited.

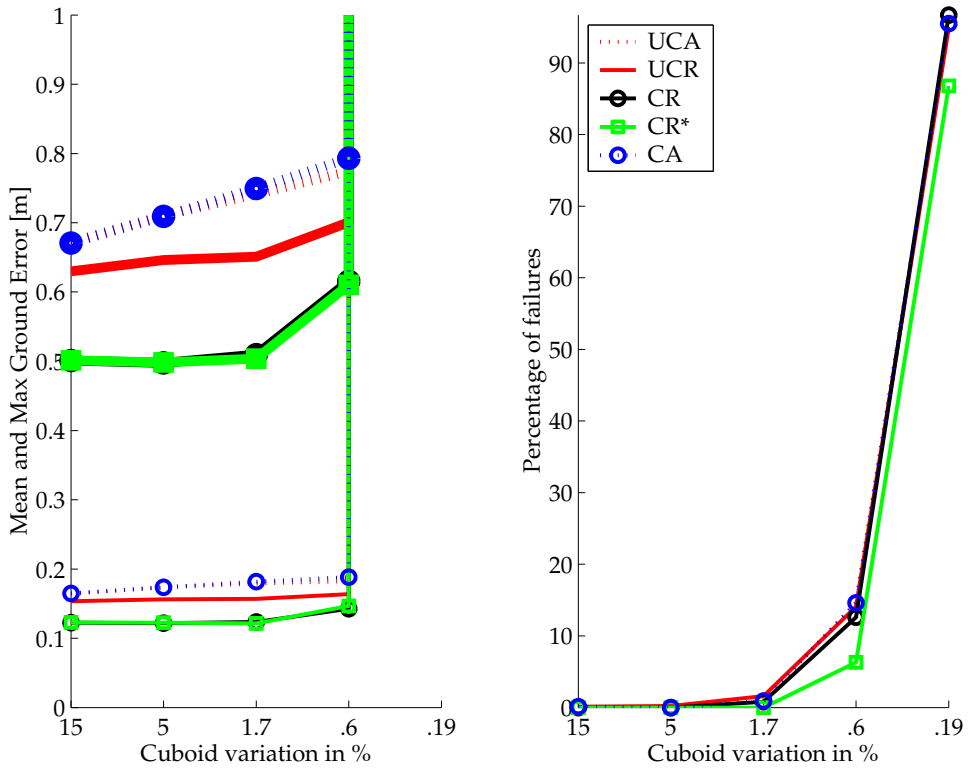

Figure 9.32: *Separation of the mean ground errors in planar component (thin line) and height component (thick line); 1000 realizations, Configuration: 'Air1', 10 points, 1 pixel noise, camera distance: 1500 m, threshold for bad solutions: 3.5 m.*

Conclusion. Summing it up, if the image geometry is not too bad and at least 10 corresponding points are used, then the simple direct linear method 'UCA' for determining the trifocal tensor, which neglects the internal constraints and minimizes reprojection error, practically performs equal to the constrained method 'CR', which minimizes reprojection error. Consequently, if the trifocal tensor is only computed to obtain approximations for the orientation parameters of a bundle block adjustment, the direct linear solution 'UCA' is usually sufficient. However, if a general routine for providing such initial values is requested, which does not have any restrictions on the image geometry and also handles weak configurations, then the Gauß-Helmert model 'CR' minimizing reprojection error must be used.

Chapter 10

Conclusion

The topic of this thesis is the **trifocal tensor**, which describes the relative orientation (or epipolar geometry) of *three uncalibrated images*. It is a homogenous valence-(1,2) tensor, which means that it can be represented as a $3\times3\times3$ cube of number. The trifocal tensor is of particular interest because it can be determined *linearly* from corresponding points or lines in the three images. Consequently it provides *a tool to determine the relative orientation of three images without requiring approximate values*. This relative orientation can be further used as initial values in a subsequent bundle block adjustment.

Each triple of points provides 4 linear equations, called *trilinearities*, in the tensor elements and each corresponding triple of lines provides 2 such linear equations; therefore at least 7 points, 13 lines or a proper combination are needed for a direct linear solution of the trifocal tensor. To obtain a unique solution it is important that the corresponding image features do not arise from a critical configuration. The probability of running into such configurations is much smaller than for two images and the fundamental matrix. However, if all corresponding image features arise from a common plane the direct solution fails. This is due to the fact that the images are considered *uncalibrated*.

10.1 Summary

Although the point and line trilinearities allow a simple direct linear solution of the trifocal tensor, there are, however, some drawbacks associated with this linear solution. The trifocal tensor is made up of 27 elements, but it only has 18 degrees of freedom. Consequently its elements have to satisfy 8 internal constraints, besides the fixing of the tensor's scale, to represent a *valid* trifocal tensor. These constraints are in generally not satisfied by the direct linear solution. Another drawback is that the direct linear solution does not minimize the errors in the original point and line measurements (so-called *reprojection error*) but some other quantities (so-called *algebraic error*).

Because of these drawbacks, errors in the tensor elements will be induced and therefore the obtained relative orientation may be not sufficient to initialize a subsequent bundle block adjustment. These drawbacks can be prevented by determining a valid trifocal tensor by minimizing reprojection error. The so-called *Gauss-Helmert model* provides a general environment

for such constrained adjustment tasks. The determination of a valid trifocal tensor in the Gauß-Helmert model can be done basically in two ways: By parameterizing the tensor using its 27 elements and introducing the 8 internal constraints together with the scale fixing, or by using an alternative parameterization for the trifocal tensor that has 18 degrees of freedom.

Several sets of constraints and alternative parameterizations have been proposed over the years and we reviewed the most important ones in the sections 7.6 on page 96 and 8.3 on page 116 respectively. In this thesis we also derived two new sets of constraints and gave a simple geometric interpretation for them. We also found a new alternative parameterization for the trifocal tensor.

For computing a valid trifocal tensor by minimizing reprojection error in the Gauß-Helmert model it is important to have a consistent representation for the trifocal tensor. The parameterization using the projection matrices proposed in [Hartley 1994a] and reviewed in section 8.3.1 on page 117 is the simplest way for such a consistent representation. This parameterization is applicable as long as (i) not all three projection centers coincide and (ii) the first projection center is different from the other two.

Empirical investigations in section 9 on page 135 showed the difference in computing the trifocal tensor by minimizing algebraic error or reprojection error with or without considering the internal constraints. Further the *minimum thickness* of the object points was investigated; i.e. what is the minimum deviation of the object points from a common plane, so that the computation is still possible?

The findings can be summarized in the following way: For image configurations with strong geometry, like in the case of convergent terrestrial images or in the case of aerial images, the direct linear solution for the trifocal tensor and the valid tensor from the Gauß-Helmert model minimizing reprojection error are practically the same. Concerning the minimum thickness unexpected small values were found already for 10 point correspondences. If the number of points increases, then the minimum thickness gets even smaller. For the mentioned configurations and 15 point correspondences the computation of the trifocal tensor is *still successful* for a minimum thickness of the object points of about 1% of the camera distance (normal angle camera with assumed noise in the images of 1 pixel).

For image configurations with weak geometry, like collinear projection centers with congruent viewing directions, much more point correspondences are required and the direct linear solution fails more often for small thicknesses of the object points than the valid solution from the Gauß-Helmert model minimizing reprojection error.

From these empirical investigations we come to the conclusion that minimizing reprojection error and considering the internal constraints is actually not necessary if the image configuration has a strong geometry – especially if one is only interested in initial values for a subsequent bundle block adjustment. However, if a general tool for providing such initial values is requested, which does not have any restrictions on the image geometry and handles also weak configurations, then the Gauß-Helmert model minimizing reprojection error must be used.

If $n > 3$ images are given, then many unrelated trifocal tensors are computed. Before all n images can be used simultaneously, the individual basis vectors and rotation matrices derived from the unrelated tensors must be transformed into one common system; see section 7.4.6 on page 94.

10.2 Contribution of this thesis

The contribution of this thesis can be divided into a theoretical part, concerning geometry and constraints of the trifocal tensor, a practical part, concerning the computation of the trifocal tensor, and an empirical part, concerning the necessity of computing a valid trifocal tensor by minimizing reprojection error as opposed to the direct linear solution.

Geometry and constraints of the trifocal tensor. Two new minimal sets of constraints were derived in section 7.6.3 on page 99. A set of constraints is termed *minimal* if it is made up of exactly 8 constraints.

For deriving these constraints we reviewed thoroughly the properties of the trifocal tensor by means of its tensorial slices, the matrices I_i, J_j and K_k; see section 7.2 on page 66.

Using the so-called *principal rays*, which are the projection rays of the canonical points of an image, we found a simple geometric interpretation for the new set of constraints: *The eight constraints of the trifocal tensor guarantee that the mappings of the principal rays of each image into the other two images are straight and concurrent lines.*

Further, we discovered a *new parameterization*, which uses 20 parameters; see section 8.3.4 on page 123. Therefore 2 constraints, which fix two scaling ambiguities, need to be considered in the Gauß-Helmert model. From this parameterization also a very simple *minimal parameterization* can be derived, which has exactly 18 parameters. These 18 parameters have a simple geometric interpretation; see proposition 8.1 on page 124.

Minimal parameterizations have the advantage that their implementation requires no additional constraints, however, generally more than one minimal parameterization are required to represent all possible trifocal tensors. In case of the new minimal parameterizations 81 maps from the set of all possible trifocal tensors to the new minimal parameterization exist. This is, however, still less than the number of maps for the minimal parameterization of [Papadopoulo and Faugeras 1998], which is 59.049.

Computation of the trifocal tensor. Using the trilinearities of points and lines in section 7.3 on page 77 the 27 elements of the trifocal tensor can be computed linearly. This direct linear solution, however, has the drawbacks, that the internal constraints are not satisfied and that algebraic error is minimized instead of reprojection error. These drawbacks can be avoided if a valid trifocal tensor is computed in the Gauß-Helmert model by minimizing reprojection error. This can be done basically in two ways: By parameterizing the tensor using its 27 elements and introducing the 8 internal constraints together with the scale fixing; or by using an alternative, minimal or non-minimal, parameterization for the trifocal tensor that has 18 degrees of freedom.

However, the method of introducing the 8 internal constraints is not recommended, because all sets of constraints reviewed in section 7.6 on page 96, including the new ones, suffer from one of the following reasons; cf. section 8.3.5 on page 128: The constraints are too complex to be implemented in the Gauß-Helmert model, they are not minimal (i.e. > 8) or they are not valid for all image configurations. Also the minimal parameterizations reviewed in section 8.3 on page 116 can not be recommended for realization in the Gauß-Helmert model due to their complexity, which is primarily caused by the many case differentiations involved.

Actually, only the new parameterization of section 8.3.4 on page 123, using 20 parameters and 2 constraints, and the parameterization using the projection matrices of section 8.3.1 on page 117, using 24 parameters and 6 constraints, are relatively simple to be realized in the Gauß-Helmert model. The benefit of the parameterization using the projection matrices is that there is only one map from all trifocal tensors to this parameterization, whereas for the new parameterization there are three maps. On the other hand the parameterization using the projection matrices requires 6 constraints, whereas the new parameterization needs only 2 constraints.

However, since the simplicity of the implementation, i.e. the number of maps, is more important than the number of required constraints, *the parameterization using the projection matrices is the recommended way to represent a valid trifocal tensor in the Gauß-Helmert model.*

Empirical studies. In chapter 9 on page 135 the trifocal tensor is determined by various methods for five different image configurations and for different numbers of corresponding points. Three image configurations, 'Tetra' (the three projection centers and the object center from a tetrahedron), 'Air1' and 'Air2' (images from one and two flight strips), represent stable image geometries with intersection angles of about 90° and 17° respectively. The other two configurations, 'Street1' and 'Street2' (three collinear projection centers with congruent viewing directions), have a rather unstable geometry with intersection angles of about 3°.

These examples are based on synthetic data and shall demonstrate the differences between *minimizing algebraic and reprojection error*, the effects of considering or neglecting the *internal constraints* of the tensor, and the impact of *critical configurations*; i.e. how close must points lie to the same plane so that the computation fails?

The results from these investigations can be summarized in the following way:

- The differences in the results of the various computation methods vanish for a particular image configuration if the *number of point correspondences increases*, and for a particular number of point correspondence if the *stability of the image geometry increases*. Therefore, from 10 points onwards for 'Tetra', 'Air1' and 'Air2', and from 25 points onwards for 'Street1' and 'Street2', all computation methods return practically the same result; i.e. then *the direct linear solution is equivalent to the constrained computation minimizing reprojection error*.

- For small numbers of point correspondences, ≤ 8 for the stable configurations 'Tetra', 'Air1' and 'Air2' and for the unstable configurations 'Street1' and 'Street2' in general, the unconstrained minimization of reprojection error generally performed better than any algebraic (constrained or unconstrained) minimization. Therefore, *the benefit of minimizing only reprojection error, is larger than of considering only the constraints*.

- The *algebraic minimizations can produce a bias* for certain lopsided configurations, i.e. 'Street2', due to the incorrect weighting of the trilinearities.

- Concerning the critical configuration of almost planar object points we can say that for image configurations with stable geometry the minimum thickness, i.e. the minimum deviation from a common plane, is practically *independent from the computation method of the tensor*. For image configurations with unstable geometry the minimum thickness strongly depends on minimizing reprojection error and considering the internal constraints.

  Generally the minimum thickness becomes smaller if the number of point correspondences increases. Therefore, the computation of the trifocal tensor is *still successful* for the following

minimum thickness of the object points, represented in units of the camera distance (normal angle camera with assumed noise in the images of 1 pixel, 15 point correspondences):

- 'Tetra' (1 pixel = 9 µm): flatness of 1/100,
- 'Air1' and 'Air2' (1 pixel = 15 µm): flatness of 6/1000.

Summing it up, if the image geometry is not too bad and at least 10 corresponding points are used, then the simple direct linear method for determining the trifocal tensor, which neglects the internal constraints and minimizes reprojection error, practically performs equal to the constrained method, which minimizes reprojection error. Consequently, if the trifocal tensor is only computed to obtain approximations for the orientation parameters of a bundle block adjustment, the direct linear solution is usually sufficient. However, if a general routine for providing such initial values is requested, which does not have any restrictions on the image geometry and also handles weak configurations, then the Gauß-Helmert model minimizing reprojection error must be used.

10.3 Outlook

The following tasks seem worth further investigations:

Calibrated images. In section 8.5 on page 132 the critical configurations for determining the trifocal tensor are reviewed. If the correspondences arise from points lying on certain curves in space or in a plane then the trifocal tensor can not be determined uniquely. The chances of running into a configuration involving those critical curves is rather small, however, point correspondences from a common plane occur very often, especially if man-made objects are involved. The examples from section 9 showed that the minimum thickness, required to successfully determine the trifocal tensor, depends on the distance between camera and object, and the noise in the image. Therefore the problem may in practise be solved by using larger image scales, however, a certain risk remains, because often the images are taken without considering these critical planarities or approaching the object is restricted by environmental conditions.

Planar object points constitute a problem because the images are regarded as *uncalibrated* when computing the trifocal tensor. If the the interior orientation of the images is known, then this problem may be solved in general. Empirically we found that the system of the direct linear solution has rank-21 instead of rank-26. Consequently by applying the internal constraints on this 5-parameter family of direct solutions a certain countable number of valid trifocal tensors will result and if we further consider the known interior orientation parameters, the number of valid trifocal tensors fitting to this interior orientation may be broken down to a single solution.

Empirical investigations. From the examples in section 9 we derived formulas for the ground errors and the minimum thickness for certain image configurations depending on the noise level in the images. These formulas are backed only by two different noise levels, which were considered in the examples, and only one principal distance is used. Therefore an approach, considering different noise levels, various principal distances and other image configurations, will result in formulas with a more general validity.

Algebraic error. We also saw during the examples, that for geometric stable configurations using a certain number of point correspondences the results obtained by minimizing algebraic error are very similar to the results of minimizing reprojection error. On the other hand we also saw that for certain lopsided image configurations the minimizing of algebraic error produces a bias in the results. A more thorough investigation of the relation ship between these two kinds of minimizations may give deeper insights into the behavior of minimizing algebraic error during the determination the trifocal tensor.

Lines. One benefit of dealing with the trifocal tensor is its ability to handle points and lines at the same time. During the examples in section 9 we only consider point correspondences. Also in the literature primarily points are used in connection with the trifocal tensor. Therefore it is worth investigating the benefits and limitations of considering line correspondences with and without additional point correspondences.

Lines have the advantage of occurring frequently in images, especially in industrial scenes. They can be extracted automatically, often with higher accuracy than points; provided the lines are long enough. Line correspondences can be established even if occlusions occur in the images. However, since one single line does not constrain the image geometry as much as one single point, many more line correspondences will be necessary to achieve the same accuracy as with a certain number of points.

Distortion. The trifocal tensor so far is embedded in the framework of projective geometry, which does not include non-linear image distortions. If the trifocal tensor is only determined to derive approximations of the image orientations for a subsequent bundle block adjustment, this limitation should not be of much concern, provided the non-linear distortions are not too large and the points are well distributed. However, if an optimal estimate of the trifocal tensor is wanted, then non-linear distortion parameters have to be included in the projective framework. Some strategies can be found in the literature; e.g. [Niini 2000] or [Fitzgibbon 2001].

Appendix A

Notation

In this chapter the notation used within this thesis is summarized in detail. Quantities with fixed scale are represented in slanted notation, and homogenous quantities are represented in upright notation.

special symbols	
$\|\mathsf{A}\| = \det(\mathsf{A})$	the determinant of the matrix A
$\|\|\mathsf{A}\|\|_{\text{Frob}}$	the Frobenius norm of the matrix A; see definition B.3 on page 181
$\|\boldsymbol{x}\|$, $\|s\|$	the length of a vector $\boldsymbol{x}$ and the absolute value of a scalar s
$\mathbf{x}^*$, $\boldsymbol{\lambda}^*$	the dual of a point $\boldsymbol{x}$ and the dual of a line $\boldsymbol{\lambda}$
$\mathsf{A}^\star$	the cofactor matrix of the matrix A, which is the transposed of the adjoint; see equation (B.14) on page 179
$<\mathbf{x}, \mathbf{y}>$	the inner, or scalar, product of two vectors $\mathbf{x}$ and $\mathbf{y}$
$\lceil x \rceil$	smallest integer being equal to or larger than x
lower case Roman font	
$\mathbf{e}'_2$, $\mathbf{e}'_3$	the epipoles in image Ψ_1 with respect to the images Ψ_2 and Ψ_3 ; see relation (6.1) on page 45
$\boldsymbol{i}_i$	the ith column of the identity matrix $\boldsymbol{I}_3$; see table 3.1 on page 15
$\mathbf{x}'$, $\mathbf{x}''$, $\mathbf{x}'''$	corresponding image points in the images Ψ_1, Ψ_2 and Ψ_3
lower case Greek font	
$\boldsymbol{\varepsilon}'_2$	an epipolar line in image Ψ_1 with respect to image Ψ_2; see section 6.1.1 on page 45
$\boldsymbol{\lambda}'$, $\boldsymbol{\lambda}''$, $\boldsymbol{\lambda}'''$	corresponding image lines in the images Ψ_1, Ψ_2 and Ψ_3
κ	the eigenvalue of a standard or general eigenvalue problem
$\boldsymbol{\pi}_\infty$	the line at infinity; see section 3.2 on page 12
$\boldsymbol{\rho}'_{31}$, $\boldsymbol{\rho}'_{32}$, $\boldsymbol{\rho}'_{33}$	maps of the principal rays $\boldsymbol{\mathcal{R}}'''_1$, $\boldsymbol{\mathcal{R}}'''_2$ and $\boldsymbol{\mathcal{R}}'''_3$ of image Ψ_3 into image Ψ_1; see proposition 6.4 on page 47

upper case Roman or Sans Serif font	
A	$\mathsf{M}_2\mathsf{M}_1{}^{-1}$; see equation (7.4) on page 63
A	the Jacobian-matrix of the condition equations, especially in the Gauß-Helmert model, with respect to the unknown parameters; see section 8.2 on page 113
B	$\mathsf{M}_3\mathsf{M}_1{}^{-1}$; see equation (7.4) on page 63
B	the Jacobian-matrix of the condition equations in the Gauß-Helmert model with respect to the observations; see section 8.2 on page 113
C_1	the calibration matrix of image Ψ_1; see equation (5.7) on page 38
C	the Jacobian-matrix of the constraints in the Gauß-Helmert model with respect to the unknown parameters; see section 8.2 on page 113
E_{12}	the essential matrix from image Ψ_2 to image Ψ_1; see equation (6.11) on page 49
F_{12}	the fundamental matrix from image Ψ_2 to image Ψ_1; see theorem 6.1 on page 46
G	a general correlation matrix; see definition 3.3 on page 22
$\mathsf{G}_{\mathcal{L}}$	the dual correlation induced by the 3D plane $\boldsymbol{\mathcal{L}}$; see proposition 6.10 on page 57
H	a general homography matrix; see definition 3.2 on page 21
H_{Ω}	the homography induced by the 3D plane $\boldsymbol{\Omega}$; see proposition 6.7 on page 53
H_{∞}	the homography induced by the plane at infinity $\boldsymbol{\Pi}_{\infty}$; see equation (6.17) on page 56
I_i	the ith tensorial slice in direction i – a dual line correlation from image Ψ_2 to image Ψ_3; see proposition 7.1 on page 67
I_n	the identity matrix of dimension n
$\boldsymbol{I}_i$	the ith column of the identity matrix I_4
J_j	the jth tensorial slice in direction j – a point homography from image Ψ_1 to image Ψ_3; see proposition 7.5 on page 71
K_k	the kth tensorial slice in direction k – a point homography from image Ψ_1 to image Ψ_2; see proposition 7.9 on page 74
M_1	the leftmost 3×3 sub-matrix of the point projection matrix P_1; see equation (5.9) on page 39
P_1	the point projection matrix of image Ψ_1; see equation (5.6) on page 38
$\mathbb{P}^1$, $\mathbb{P}^2$, $\mathbb{P}^3$, $\mathbb{P}^n$	the projective spaces of dimension 1, 2, 3 and n
$\mathbb{P}^{2*}$	the dual projective spaces of dimension 2
Q_1	the line projection matrix of image Ψ_1; see proposition 5.1 on page 40
R_1	the rotation matrix of image Ψ_1; see section 5.2 on page 37
$\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^4$, $\mathbb{R}^n$	the Euclidian spaces of dimension 2, 3, 4 and n

upper case Roman or Sans Serif font	
$\mathsf{S}(\cdot)$	the axiator, i.e. the operator that returns the cross-product matrix; see equation (B.8) on page 179
$\mathsf{S}^{red}(\cdot)$	the axiator reduced by one row; see section 3.6 on page 27
S	the diagonal matrix of singular values; cf. section B.2 on page 178
U	the left orthogonal matrix of singular value decomposition
V	the right orthogonal matrix of singular value decomposition
$\mathbf{Z}_1$	the projection center of image Ψ_1; always represented as a Euclidian vector

upper case Greek font	
$\Gamma(\cdot)$	the operator that returns a 4×4 matrix, the Plücker matrix, needed for joining and intersecting 3D entities; see section 3.5 on page 23
$\mathbf{\Sigma}(\boldsymbol{\lambda}')$	the projection plane of image line $\boldsymbol{\lambda}'$ in image Ψ_1; see relation (5.20) on page 42
$\Pi(\cdot)$	the operator that returns a 6×4 matrix needed for joining and intersecting 3D entities; see section 3.5 on page 23
$\mathbf{\Pi}'_1, \mathbf{\Pi}'_2, \mathbf{\Pi}'_3$	the principal planes of image Ψ_1
$\mathbf{\Pi}_\infty$	the plane at infinity
Ψ_1	the first image; see relation (5.13) on page 39
$\mathbf{\Omega}$	a general plane in 3D space

upper case Calligraphic font	
$\mathcal{I}_i$	the ith column of the identity matrix I_6; see table 3.3 on page 21
$\mathcal{L}$	a general line in 3D space
$\mathcal{N}(\mathsf{A})$	the nullspace of a matrix A
$\mathcal{R}'_1, \mathcal{R}'_2, \mathcal{R}'_3$	the principal rays of image Ψ_1 see relation (5.18) on page 41
$\mathcal{R}(\mathsf{A})$	the range of a matrix A
$\mathcal{S}(\mathbf{x}')$	the projection ray of image point $\mathbf{x}'$ in image Ψ_1 see relation (5.23) on page 42

Appendix B

Matrix Calculus

This chapter briefly summarizes some concepts of matrix calculus, which are used within this thesis.

B.1 The eigenvalue problem

Let A be a square matrix of size $n \times n$. Then finding solutions for the scalars $\kappa_1, \kappa_2, \ldots$ and the vectors $x_1, x_2, \ldots$ of the equation

$$Ax = \kappa x \tag{B.1}$$

is called the (standard) **eigenvalue problem**. The scalars κ_i are called the (standard) *eigenvalues* and the vectors x_i are called the (standard) *right eigenvectors* of matrix A. The equation $x^\top A = \kappa x^\top$ has the same eigenvalues but in general different (standard) *left eigenvectors*.

The **general eigenvalue problem** is of the form

$$Ax = \kappa Bx, \tag{B.2}$$

where B is also of size $n \times n$. The scalars κ are called *general eigenvalues* and the vectors x are called the *general right* resp. *left eigenvectors*. Without any further specification the term *eigenvalue problem* always refers to the standard case.

Eigenvalues are found as the roots of the characteristic polynomial

$$\det(A - \kappa I) = 0. \tag{B.3}$$

The multiplicity a_i of a root κ_i of this polynomial is called *algebraic multiplicity* of the eigenvalue κ_i. The set of eigenvectors x that satisfies (B.1) for a particular eigenvalue κ_i is called the *eigenspace* of the eigenvalue κ_i, and its dimension $g_i = n - \operatorname{rank}(A - \kappa_i I)$ is called the *geometric multiplicity* of the eigenvalue κ_i. Both types of multiplicities are related by the following important inequality; e.g. [Überhuber 1997]

$$1 \leq g_i \leq a_i. \tag{B.4}$$

The eigenvalues of a matrix A are related to the determinant and the trace in the following way:

$$\prod \kappa_i = \det(A), \tag{B.5}$$

$$\sum \kappa_i = \operatorname{trace}(A). \tag{B.6}$$

B.2 The singular value decomposition

The *singular value decomposition* (SVD) of a real matrix A of size $m \times n$ is given by

$$A = USV^\top, \tag{B.7}$$

where U and V are orthogonal matrices of size $m \times m$ and $n \times n$ respectively, and $S = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k)$ is a diagonal matrix of size $m \times n$ and $k = \min(m, n)$. The diagonal elements satisfy the following inequality

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r = \sigma_{r+1} = \cdots = \sigma_k = 0 \quad \text{with} \quad r = \operatorname{rank}(A).$$

These diagonal elements are called *singular values* of the matrix A and are identical to the positive square-roots of the eigenvalues of $A^\top A$ and $AA^\top$. The column vectors of U are called *left singular vectors* of A and correspond to the eigenvectors of $AA^\top$, and the column vectors of V are called *right singular vectors* of A and correspond to the eigenvectors of $A^\top A$; e.g. [Golub and VanLoan 1996]. The matrices U and V and hence the singular value decomposition are unique, if all singular values have multiplicity 1.

The singular value decomposition is a very useful tool for doing linear algebra – both in a numerical and theoretical way; see e.g. [Golub and VanLoan 1996], [Überhuber 1997] or [Hartley and Zisserman 2001].

For example, the range $\mathcal{R}(A)$ and nullspace $\mathcal{N}(A)$ of a rank-r matrix A is easily obtained by

$$\begin{aligned}\mathcal{R}(A) &= \boldsymbol{u}_1 \vee \boldsymbol{u}_2 \vee \cdots \vee \boldsymbol{u}_r \\ \mathcal{N}(A) &= \boldsymbol{v}_{r+1} \vee \boldsymbol{v}_{r+2} \vee \cdots \vee \boldsymbol{v}_n,\end{aligned}$$

where $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ denotes the ith column of U and V respectively; see e.g. [Überhuber 1997].

Consequently, the singular value decomposition represents a simple way to solve homogenous systems of linear equations $A\boldsymbol{x} = 0$, which we encounter several times in this thesis. Provided the system has rank-$(n-1)$, the unique solution $\overline{\boldsymbol{x}}$ is obtained by the last column of V; i.e. $\mathcal{N}(A) = \boldsymbol{v}_n$. In case of an over-determined homogenous system (i.e. $m > n$), the same vector $\overline{\boldsymbol{x}} = \boldsymbol{v}_n$, being the left singular vector to the smallest singular value σ_n, is the solution of

$$|A\boldsymbol{x}|^2 \to \min$$

and $|A\boldsymbol{x}| = \sigma_n$. See e.g. [Hartley and Zisserman 2001] for further possible usages of the singular value decomposition concerning linear least squares problems.

Also the *pseudo inverse* or *Moore-Penrose inverse* A^+ of a *singular* matrix A can be obtained easily with the singular value decomposition:

$$A^+ = VS^+U^\top,$$

with $S^+ = \operatorname{diag}(1/\sigma_1, 1/\sigma_2, \ldots, 1/\sigma_r, 0, \ldots, 0)$. The pseudo satisfies the following conditions

$$\begin{aligned}AA^+A &= A \\ A^+AA^+ &= A^+ \\ (A^+A)^\top &= A^+A \\ (AA^+)^\top &= AA^+ .\end{aligned}$$

B.3 The cross product

The cross product of two vectors $\boldsymbol{a} = (a_1, a_2, a_3)^\top$ and $\boldsymbol{b} = (b_1, b_2, b_3)^\top$, denoted by $\boldsymbol{a} \times \boldsymbol{b}$, returns a vector $\boldsymbol{c}$ that is orthogonal to both vectors $\boldsymbol{a}$ and $\boldsymbol{b}$:

$$\boldsymbol{c} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \quad \text{with} \quad \boldsymbol{c}^\top \boldsymbol{a} = \boldsymbol{c}^\top \boldsymbol{b} = 0.$$

Using the so-called *axiator* $\mathsf{S}(\boldsymbol{a})$, this can also be written as

$$\boldsymbol{c} = \mathsf{S}(\boldsymbol{a})\boldsymbol{b},$$

with

$$\mathsf{S}(\mathbf{a}) = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \tag{B.8}$$

In other publications the axiator is represented as $[\,\mathbf{a}\,]_\times$; e.g. [Hartley and Zisserman 2001]. The following relations involving $\mathsf{S}(\boldsymbol{a})$ are very helpful.

$$\alpha\mathsf{S}(\boldsymbol{a}) + \beta\mathsf{S}(\boldsymbol{b}) = \mathsf{S}(\alpha\boldsymbol{a} + \beta\boldsymbol{b}) \tag{B.9}$$

$$\mathsf{S}(\boldsymbol{a})\mathsf{S}(\boldsymbol{b}) = \boldsymbol{b}\boldsymbol{a}^\top - \boldsymbol{a}^\top\boldsymbol{b}\boldsymbol{I} \tag{B.10}$$

$$\mathsf{S}(\boldsymbol{a}\times\boldsymbol{b}) = \mathsf{S}(\mathsf{S}(\boldsymbol{a})\boldsymbol{b}) = \boldsymbol{b}\boldsymbol{a}^\top - \boldsymbol{a}\boldsymbol{b}^\top \tag{B.11}$$

$$\mathsf{S}(M\boldsymbol{a}) = \det(M)M^{-\top}\mathsf{S}(\boldsymbol{a})M^{-1}, \tag{B.12}$$

with M regular. The general formulation of (B.12), which holds for any M, is

$$\mathsf{S}(M\boldsymbol{a})M = \mathrm{adj}(M^\top)\mathsf{S}(\boldsymbol{a}). \tag{B.13}$$

$\mathrm{adj}(M)$ is the *adjoint* of the matrix M, defined by

$$\mathrm{adj}(M)(i,j) = (-1)^{i+j}\det(M'_{ji}), \tag{B.14}$$

where M'_{ji} denotes the matrix obtained by striking out the jth row and the ith column of M. So, if M is regular, it holds $\mathrm{adj}(M) = \det(M)M^{-1}$. The relation (B.13) is adopted from [Hartley and Zisserman 2001]; the other relations are obtained by simple computations.

Here are some rules that hold for cross products involving more than two vectors, adapted from [Bronstein and Semendjajew 1991].

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) \neq (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}$$

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \mathsf{S}(\boldsymbol{a})\mathsf{S}(\boldsymbol{b})\boldsymbol{c} \tag{B.15}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = \mathsf{S}(\mathsf{S}(\boldsymbol{a})\boldsymbol{b})\boldsymbol{c} \tag{B.16}$$

$$(\boldsymbol{a} \times \boldsymbol{b})^\top(\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a}^\top\boldsymbol{c})(\boldsymbol{b}^\top\boldsymbol{d}) - (\boldsymbol{a}^\top\boldsymbol{d})(\boldsymbol{b}^\top\boldsymbol{c}) \tag{B.17}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \times (\boldsymbol{c} \times \boldsymbol{d}) = |\boldsymbol{abd}|\boldsymbol{c} - |\boldsymbol{abc}|\boldsymbol{d} = |\boldsymbol{acd}|\boldsymbol{b} - |\boldsymbol{bcd}|\boldsymbol{a} \tag{B.18}$$

$$(\boldsymbol{a} \times \boldsymbol{b})^\top\big((\boldsymbol{c} \times \boldsymbol{d}) \times (\boldsymbol{e} \times \boldsymbol{f})\big) = |\boldsymbol{abd}||\boldsymbol{cef}| - |\boldsymbol{abc}||\boldsymbol{def}| \tag{B.19}$$

Here $|\boldsymbol{abc}|$ denotes the determinant of the matrix made up of the column vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$.

B.4 The Kronecker product and the vec ()-operator

The Kronecker product and the vec ()-operator are useful for rewriting vector and matrix relations.

Definition B.1 (The Kronecker product.) *Let A by a matrix of size $m \times n$ and B by a matrix of size $p \times q$. Then the Kronecker product $A \otimes B$ defines a matrix of size $mp \times nq$ as*

$$A \otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

Definition B.2 (The vec ()-operator.) *Let $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n]$ by a matrix of size $m \times n$. Then $\mathrm{vec}\,(A)$ returns a vector of size $mn \times 1$ by stacking together the columns $\boldsymbol{a}_i$ of A:*

$$\mathrm{vec}\,(A) := \left(\boldsymbol{a}_1^\top, \boldsymbol{a}_2^\top, \ldots, \boldsymbol{a}_n^\top\right)^\top.$$

The following relations involving the Kronecker product and the vec ()-operator are very helpful; see e.g. [Schmidt and Trenkler 1998].

$$(A \otimes B)C = A \otimes C + B \otimes C \tag{B.20}$$
$$C(A \otimes B) = C \otimes A + C \otimes B \tag{B.21}$$
$$(A \otimes B)^\top = A^\top \otimes B^\top \tag{B.22}$$
$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \tag{B.23}$$
$$\text{with} \quad A : (m \times n), B : (p \times q), C : (n \times r), D : (q \times s)$$

$$\mathrm{vec}\,(\alpha A) + \mathrm{vec}\,(\beta B) = \alpha\,\mathrm{vec}\,(A) + \beta\,\mathrm{vec}\,(B) \tag{B.24}$$
$$\mathrm{vec}\,(\boldsymbol{a}) = \mathrm{vec}\left(\boldsymbol{a}^\top\right) = \boldsymbol{a} \tag{B.25}$$
$$\mathrm{vec}\left(\boldsymbol{a}\boldsymbol{b}^\top\right) = \boldsymbol{b} \otimes \boldsymbol{a} \tag{B.26}$$
$$\mathrm{vec}\,(ABC) = (C^\top \otimes A)\,\mathrm{vec}\,(B) \tag{B.27}$$

By some simple but tedious computations we can also see that the following relation holds

$$A\,\mathrm{vec}\,(B) = A T_n\,\mathrm{vec}\left(B^\top\right), \tag{B.28}$$

with A of size $m \times n^2$, B of size $n \times n$ and T_n of the following form

$$\underset{n^2 \times n^2}{T_n} = \left[\boldsymbol{i}_1, \boldsymbol{i}_{1+n}, \boldsymbol{i}_{1+2n}, \ldots, \boldsymbol{i}_{1+(n-1)n}, \boldsymbol{i}_2, \boldsymbol{i}_{2+n}, \ldots, \boldsymbol{i}_{2+(n-1)n}, \ldots, \boldsymbol{i}_n, \boldsymbol{i}_{n+n}, \ldots, \boldsymbol{i}_{n+(n-1)n}\right], \tag{B.29}$$

where $\boldsymbol{i}_i$ is the ith column of the identity matrix I_{n^2}.

B.5 Miscellaneous

At some points of this thesis we need one of the following terms or relations associated with matrices.

Proposition B.1 (A special inverse.) *The inverse of the matrix* $M = cI + \boldsymbol{a}\boldsymbol{b}^\top$ *is given by*

$$M^{-1} = \frac{1}{c}(I - \frac{1}{c + \boldsymbol{a}^\top \boldsymbol{b}}\boldsymbol{a}\boldsymbol{b}^\top),$$

provided $\boldsymbol{a}^\top \boldsymbol{b} \neq -c$ *and* $c \neq 0$*, because in these cases* M *is singular.*

Proof:
It is easy to show by multiplying M with the proposed form of M^{-1} that the result is the identity matrix I. What remains to show are the conditions for M to be singular. Of course $c \neq 0$ is required, because otherwise only the dyad $\boldsymbol{a}\boldsymbol{b}^\top$ with rank-1 remains. So we can cancel out c from our problem and consider only the $n \times n$ matrix $M' = I + \boldsymbol{a}'\boldsymbol{b}^\top$, with $\boldsymbol{a}' = 1/c\boldsymbol{a}$. From requiring $\det(M') = 0$ we see that -1 is an eigenvalue of $\boldsymbol{a}'\boldsymbol{b}^\top$. If we look therefore at the eigenvalue problem $\boldsymbol{a}'\boldsymbol{b}^\top + \kappa I$, then we see that also 0 is an eigenvalue and as the resulting matrix is the dyad $\boldsymbol{a}'\boldsymbol{b}^\top$, the rank defect and therefore the geometric multiplicity is $n - 1$. From inequality (B.4) on page 177 we see that the algebraic multiplicity has to be at least $n - 1$. So, $\boldsymbol{a}'\boldsymbol{b}^\top$ only has two different eigenvalues: $\kappa_1 = -1$ with multiplicity 1, and $\kappa_2 = 0$ with multiplicity $n - 1$.

By applying relation (B.6) on page 177 with $\sum \kappa_i = \kappa_1$ and $\text{trace}(\boldsymbol{a}'\boldsymbol{b}^\top) = \boldsymbol{a}'^\top \boldsymbol{b}$, we see that $\boldsymbol{a}'^\top \boldsymbol{b} = -1$. And consequently $\boldsymbol{a}^\top \boldsymbol{b} = -c$ is required for M to be singular.

□

Definition B.3 (Frobenius norm.) *Let* A *by a matrix of size* $m \times n$*. Then the Frobenius norm* $||A||_{\text{Frob}}$ *is defined as*

$$||A||_{\text{Frob}} = (\sum a_{ij}^2)^{1/2}.$$

Appendix C

The Algebraic Derivation of the Multiple View Tensors

In [Hartley 1995b] and [Heyden 2000] the following approach of deriving the multiple view tensors (i.e. fundamental matrix, trifocal tensor and quadrifocal tensor) is presented. From this approach it is easy to see that such tensor for the relative orientation of more the four images is not possible.

The general idea. The map $\mathbf{x}$ of a 3D point $\mathbf{X}$ into an image can be represented using the point projection matrix P as

$$\mu\mathbf{x} = \mathsf{P}\mathbf{X}, \tag{C.1}$$

with a scalar μ replacing the scale ambiguity in the notation of section 5.2. Given m images with mappings $\mathbf{x}'$, $\mathbf{x}''$, $\mathbf{x}'''$, $\ldots$, $\mathbf{x}^{(m)}$ of the point $\mathbf{X}$, we can look for *relations* that have to hold between *all* corresponding image points $\mathbf{x}', \ldots, \mathbf{x}^{(m)}$. For finding such relations we consider the following system of equations

$$\begin{bmatrix} \mathsf{P}_1 & \mathbf{x}' & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{x}'' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathsf{P}_3 & \mathbf{0} & \mathbf{0} & \mathbf{x}''' & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathsf{P}_m & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x}^{(m)} \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -\mu_1 \\ -\mu_2 \\ -\mu_3 \\ \vdots \\ -\mu_m \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}. \tag{C.2}$$

In a short way we can represent this system as

$$\mathsf{N}\mathbf{u} = \mathbf{0},$$

where N is a $3m \times (4+m)$ matrix and $\mathbf{u}$ is a $(4+m)$ vector. A necessary and sufficient condition for the homogenous system (C.2) to have a non-trivial solution is that

$$\operatorname{rank}(\mathsf{N}) < m + 4. \tag{C.3}$$

or in other words any $(m+4) \times (m+4)$ determinant of N has to be zero. Obviously these determinant conditions are good candidates to derive relations between all corresponding image points $\mathbf{x}', \ldots, \mathbf{x}^{(m)}$. Further, since the image points appear in separate columns in N, these relations will be *multi-linear* in these image points.

Two images. At first we consider the case of two images ($m = 2$). The system (C.2) turns therefore into

$$\begin{bmatrix} \mathsf{P}_1 & \mathbf{x}' & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{x}'' \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -\mu_1 \\ -\mu_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

This is a 6×6 system and therefore only one determinant of size $(m + 4)$ has to be zero. By denoting the rows of P_1 with $\mathbf{A}^i$ and the rows of P_2 with $\mathbf{B}^i$, this determinant condition looks the following

$$\det \begin{bmatrix} \mathbf{A}^1 & x'^1 & 0 \\ \mathbf{A}^2 & x'^2 & 0 \\ \mathbf{A}^3 & x'^3 & 0 \\ \mathbf{B}^1 & 0 & x''^1 \\ \mathbf{B}^2 & 0 & x''^2 \\ \mathbf{B}^3 & 0 & x''^3 \end{bmatrix} = 0.$$

This determinant condition already corresponds to the epipolar constraint 6.2. Using the fully antisymmetric valence-3 tensor ϵ_{ijk} (see section 4) the expansion of this determinant gives the following bilinear relation (observe the summation convention)

$$\frac{1}{4}\epsilon_{ief}x'^i\epsilon_{jpq}x''^j \det \begin{bmatrix} \mathbf{A}^e \\ \mathbf{A}^f \\ \mathbf{B}^p \\ \mathbf{B}^q \end{bmatrix} = 0,$$

with each index running from 1 to 3. The quantities not depending on the image coordinates $\mathbf{x}'$ and $\mathbf{x}''$ can be stuck together to form a matrix F_{ij} as

$$F_{ij} = \frac{1}{4}\epsilon_{ief}\epsilon_{jpq} \det \begin{bmatrix} \mathbf{A}^e \\ \mathbf{A}^f \\ \mathbf{B}^p \\ \mathbf{B}^q \end{bmatrix}.$$

This bifocal tensor corresponds to the *fundamental matrix* F_{12} in theorem 6.1. The epipolar constraint in tensor notation is then represented by

$$F_{ij}x'^i x''^j = 0.$$

Three images. In the case of three images $(m = 3)$ the system (C.2) turns into

$$\begin{bmatrix} \mathsf{P}_1 & \mathbf{x}' & \mathbf{0} & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{x}'' & \mathbf{0} \\ \mathsf{P}_3 & \mathbf{0} & \mathbf{0} & \mathbf{x}''' \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -\mu_1 \\ -\mu_2 \\ -\mu_3 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

This is a 9×7 system and since $m + 4 = 7$ any 7×7 determinant has to be zero. Basically we have two possibilities to choose 7 rows: (i) 3 rows from one camera, 3 rows from another camera and 1 row from the last camera, or (ii) 3 rows from one camera and 2 rows from each of the

other two cameras. The first kind of determinants, however, will not yield a relation between all three image points. This can be seen by considering the following determinant condition

$$\det \begin{bmatrix} \mathbf{A}^1 & x'^1 & 0 & 0 \\ \mathbf{A}^2 & x'^2 & 0 & 0 \\ \mathbf{A}^3 & x'^3 & 0 & 0 \\ \mathbf{B}^1 & 0 & x''^1 & 0 \\ \mathbf{B}^2 & 0 & x''^2 & 0 \\ \mathbf{B}^3 & 0 & x''^3 & 0 \\ \mathbf{C}^1 & 0 & 0 & x'''^1 \end{bmatrix} = 0,$$

where the rows of P_3 are identified with $\mathbf{C}^i$. Expanding this determinant by the last column gives

$$x'''^1 \det \begin{bmatrix} \mathbf{A}^1 & x'^1 & 0 \\ \mathbf{A}^2 & x'^2 & 0 \\ \mathbf{A}^3 & x'^3 & 0 \\ \mathbf{B}^1 & 0 & x''^1 \\ \mathbf{B}^2 & 0 & x''^2 \\ \mathbf{B}^3 & 0 & x''^3 \end{bmatrix} = 0,$$

which is independent on the third camera. This is in contrast to what we aimed for – conditions depending on *all* corresponding image points. Consequently we only can consider determinants of the second type, where 3 rows are taken only from one camera.

If we use all rows from the first camera and the rows (p, q) and (r, s) from the second and third camera respectively, we get the following determinant condition

$$\det \begin{bmatrix} \mathbf{A}^1 & x'^1 & 0 & 0 \\ \mathbf{A}^2 & x'^2 & 0 & 0 \\ \mathbf{A}^3 & x'^3 & 0 & 0 \\ \mathbf{B}^{(p)} & 0 & x''^{(p)} & 0 \\ \mathbf{B}^{(q)} & 0 & x''^{(q)} & 0 \\ \mathbf{C}^{(r)} & 0 & 0 & x'''^{(r)} \\ \mathbf{C}^{(s)} & 0 & 0 & x'''^{(s)} \end{bmatrix} = 0,$$

with super-scripts in brackets indicating *fixed* indices. Since we have 3 choices for leaving one of the **B**-rows and also 3 choices for leaving one of the **C**-rows, we get actually 9 different conditions involving the corresponding points in three images. Of course, these conditions correspond to the 9 point-trilinearities (7.29) on page 79.

Once again we use the fully antisymmetric valence-3 tensor ϵ_{ijk} and represent these 9 conditions as the following trilinear relations

$$\frac{1}{2}\epsilon_{ief} x'^i \epsilon_{jpq} x''^p \epsilon_{krs} x'''^r \det \begin{bmatrix} \mathbf{A}^e \\ \mathbf{A}^f \\ \mathbf{B}^j \\ \mathbf{C}^k \end{bmatrix} = 0_{qs}, \tag{C.4}$$

with all indices running from 1 to 3. By combining the quantities which are not related to the image coordinates x'^i, x''^p and x'''^r and common to all 9 equations, i.e. do not depend on q and

s, we get the following valence (1,2) tensor T_i^{jk}

$$T_i^{jk} = \frac{1}{2}\epsilon_{ief} \det \begin{bmatrix} \mathbf{A}^e \\ \mathbf{A}^f \\ \mathbf{B}^j \\ \mathbf{C}^k \end{bmatrix} . \tag{C.5}$$

Note, that there is a summation over e and f. This tensor corresponds to the *trifocal tensor*[1] in definition 7.1. Consequently, all the transfer relations and trilinearities can also be represented in tensor notation – see table C.1.

transfer relations			
entities	equations		
$\{\boldsymbol{\lambda}'', \boldsymbol{\lambda}'''\} \rightarrow \boldsymbol{\lambda}'$	$\boldsymbol{\lambda}'_i \sim \boldsymbol{\lambda}''_j \boldsymbol{\lambda}'''_k T_i^{jk}$		
$\{\mathbf{x}', \boldsymbol{\lambda}'''\} \rightarrow \mathbf{x}''$	$\mathbf{x}''^j \sim \mathbf{x}'^i \boldsymbol{\lambda}'''_k T_i^{jk}$		
$\{\mathbf{x}', \boldsymbol{\lambda}''\} \rightarrow \mathbf{x}'''$	$\mathbf{x}'''^k \sim \mathbf{x}'^i \boldsymbol{\lambda}''_j T_i^{jk}$		
trilinearities			
entities	equations	m	n
$\boldsymbol{\lambda}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''$	$\epsilon^{abi}\boldsymbol{\lambda}'_b \boldsymbol{\lambda}''_j \boldsymbol{\lambda}'''_k T_i^{jk} = 0^a$	3	2
$\mathbf{x}', \boldsymbol{\lambda}'', \boldsymbol{\lambda}'''$	$\mathbf{x}'^i \boldsymbol{\lambda}''_j \boldsymbol{\lambda}'''_k T_i^{jk} = 0$	1	1
$\mathbf{x}', \mathbf{x}'', \boldsymbol{\lambda}'''$	$\mathbf{x}'^i \epsilon_{abj} \mathbf{x}''^b \boldsymbol{\lambda}'''_k T_i^{jk} = 0_a$	3	2
$\mathbf{x}', \boldsymbol{\lambda}'', \mathbf{x}'''$	$\mathbf{x}'^i \epsilon_{abj} \mathbf{x}'''^b \boldsymbol{\lambda}'''_k T_i^{jk} = 0_b$	3	2
$\mathbf{x}', \mathbf{x}'', \mathbf{x}'''$	$\mathbf{x}'^i \epsilon_{abj} \mathbf{x}''^b \epsilon_{cdk} \mathbf{x}'''^d T_i^{jk} = 0_{ab}$	9	4

Table C.1: *The transfer relations and trilinearities associated with the trifocal tensor in tensor notation; cf. table 7.5 on page 80. The last two columns for the trilinearities give the number m of all equations and the number n of the linearly independent equations within one group of trilinearities.*

Due to inequality (C.3) we have to consider the minors of size $(m + 4)$ of the system (C.2) for m images and as we just saw from each image at least 2 rows must be included, otherwise we will not obtain a relation between all corresponding image points. Therefore the following inequality must hold

$$m + 4 \geq 2m \qquad \rightarrow \qquad m \leq 4.$$

Consequently we are limited to the maximum of four images, when trying to find multi-linear relations for corresponding points in all images.

[1] In contrast to the geometric based derivation of the trifocal tensor in section 7.1, no particular order of the three indices i, j and k results from this algebraic derivation, therefore the tensor can be addressed as T_i^{jk}; and it holds $T_i^{jk} = mT_{\cdot\cdot i}^{kj}$ for some non-zero scalar m.

Four images. In the the last possible case of four images $(m = 4)$ the system (C.2) turns into

$$\begin{bmatrix} \mathsf{P}_1 & \mathbf{x}' & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathsf{P}_2 & \mathbf{0} & \mathbf{x}'' & \mathbf{0} & \mathbf{0} \\ \mathsf{P}_3 & \mathbf{0} & \mathbf{0} & \mathbf{x}''' & \mathbf{0} \\ \mathsf{P}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{x}'''' \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -\mu_1 \\ -\mu_2 \\ -\mu_3 \\ -\mu_4 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

This is a 12×8 system and since $m + 4 = 8$ any 8×8 determinant has to be zero. Since we have to choose at least two rows from each image, we arrive at the following determinant condition

$$\det \begin{bmatrix} \mathbf{A}^i & x'^i & 0 & 0 & 0 \\ \mathbf{A}^p & x'^p & 0 & 0 & 0 \\ \mathbf{B}^j & 0 & x''^j & 0 & 0 \\ \mathbf{B}^q & 0 & x''^q & 0 & 0 \\ \mathbf{C}^k & 0 & 0 & x'''^k & 0 \\ \mathbf{C}^r & 0 & 0 & x'''^r & 0 \\ \mathbf{D}^l & 0 & 0 & 0 & x''''^l \\ \mathbf{D}^s & 0 & 0 & 0 & x''''^s \end{bmatrix} = 0,$$

where (i, p), (j, q), (k, r) and (l, s) denotes the rows chosen from the first, second, third and fourth image respectively. The rows of the fourth image are represented as $\mathbf{D}^l$. Since we have 3 choices for dropping one row from each projection matrix, we actually get 81 different conditions involving the corresponding points in four images.

Using the fully antisymmetric valence-3 tensor ϵ_{ijk} we represent these 81 determinant conditions as quadrilinear relations

$$\epsilon_{ipt} x'^p \epsilon_{jqu} x''^q \epsilon_{krv} x'''^r \epsilon_{lsw} x''''^s \det \begin{bmatrix} \mathbf{A}^i \\ \mathbf{B}^j \\ \mathbf{C}^k \\ \mathbf{D}^l \end{bmatrix} = 0_{tuvw}, \tag{C.6}$$

with each index running from 1 to 3. By combining the quantities which are not related to the image coordinates x'^p, x''^q, x'''^r and x''''^s, and common to all 81 equations, i.e. do not depend on t, u, v and w, we get the following valence (0,4) tensor Q^{ijkl}

$$Q^{ijkl} = \det \begin{bmatrix} \mathbf{A}^i \\ \mathbf{B}^j \\ \mathbf{C}^k \\ \mathbf{D}^l \end{bmatrix}.$$

This tensor is called *quadrifocal tensor*, which is made up of $3^4 = 81$ elements. The 81 relations (C.6) involving four corresponding image points are called *quadrilinearities*, which therefore can be written as

$$\epsilon_{ipt} x'^p \epsilon_{jqu} x''^q \epsilon_{krv} x'''^r \epsilon_{lsw} x''''^s Q^{ijkl} = 0_{tuvw}, \tag{C.7}$$

Of these 81 relations, however, only 16 are linearly independent. This can be see from the fact, that in (C.7) each image point appears in combination with an ϵ-tensor; e.g. $\epsilon_{ipt} x'^p$. This corresponds to the axiator $\mathsf{S}(\mathbf{x}')$, which only has rank-2. If we consider one of these 81 equations by

fixing the indices t, u, v and w, we use from each axiator one column, which corresponds to one line through the respective image point. Obviously only two of the three lines through each point are linearly independent, therefore only $2^4 = 16$ independent equations are obtained.

One needs nevertheless at least 6 (instead of 5) point quadruples to determine the quadrifocal tensor, because there exists one linear dependency between the 32 quadrilinear relations obtained from two different point quadruples in four images; [Heyden 2000].

The monofocal tensor. In [Heyden 2000] another determinant is considered, which is made up of all three rows $\mathbf{A}^1$, $\mathbf{A}^2$ and $\mathbf{A}^3$ of the first point projection matrix P_1 and one row $\mathbf{B}^i$ from the second point projection matrix P_2

$$\det \begin{bmatrix} \mathbf{A}^1 \\ \mathbf{A}^2 \\ \mathbf{A}^3 \\ \mathbf{B}^i \end{bmatrix} .$$

This determinant does not depend on any image coordinates (and will be non-zero in general). As it is shown in [Heyden 2000] the three possible determinants of this kind make up the epipole $\mathbf{e}''_1$, i.e. the map of the first projection center into the second image, and is also called *monofocal tensor*

$$e''^i_1 = \det \begin{bmatrix} \mathbf{A}^1 \\ \mathbf{A}^2 \\ \mathbf{A}^3 \\ \mathbf{B}^i \end{bmatrix} .$$

Appendix D

Least Squares Adjustment

In the following sections the basics of least squares adjustment will be presented to the reader. Although it was tried to formulate this chapter as general as possible, it is beyond the scope of this thesis to cover all aspects of adjustment (especially the statistics part) and the interested reader is recommend to the text books [Mikhail 1976] and [Koch 1999], which are also the basis for this chapter.

D.1 Introduction

Let us consider some sort of *natural object*, which can have a *static* (time invariant) or *dynamic* (time dependent) behavior. An example of a static object would be a building, whereas the Earth's atmosphere would be a dynamic object. Now, we would like to determine and represent such a natural object. It is evident, that we can not describe the natural object per se but only an approximation of it using an *object model*. Such an object model can be very complex (e.g. the model of a building considers its doors, windows, and even the chimneys) or rather simple (e.g. the building is represented only as a cuboid).

The object model itself is represented by a set of *parameters* $\boldsymbol{x}_{Object}$; e.g. a cuboid model of a building could be represented by the length of the three edges or by the eight corner points. The parameters can be subdivided into *known* (or *constants*) and *unknown* parameters. Of course our aim is to determine the unknown parameters and thus complete our object model. This task can be solved by performing *measurements* (or *observations*) $\boldsymbol{m}$ with a suitable device (e.g. a tape or a camera). Sometimes we can directly observe an unknown parameter with a device (like the distance between to points with a tape), but usually we can only measure quantities, which are functionally related to the unknowns in a certain way (e.g. points, which we can observe in an image of a building are related to the object points through a central projection).

The observations $\boldsymbol{m}$, which are the basis for our work, however, are affected by some sort of errors, which can be seen from the fact that repeated observations of the same quantity return (slightly) different values each time. Thus, if we would make only as many observations as we need for a unique determination of our unknown parameters, the errors of the observations would affect the parameters in an uncontrolled way. So, to limit the influence of the errors on the result and to assess the effect of the remaining errors, many more observations than actually necessary are made. Now the question arises, how to determine the unknowns in a *best* way given this redundant information.

D.2 The Mathematical Model

Before we can determine the parameters of our object model, we have to set up the so-called *mathematical model*, which itself consists of the *functional model* and the *stochastic model*.

The *functional model* describes the relations $\boldsymbol{F}(\boldsymbol{m}, \boldsymbol{x}_{Object}, \boldsymbol{x}_{Instrument}) = \boldsymbol{0}$ between the object parameters $\boldsymbol{x}_{Object}$ and the observations $\boldsymbol{m}$ and introduces the additional (un)known parameters $\boldsymbol{x}_{Instrument}$. These additional parameters are required for establishing the functional relations $\boldsymbol{F}$, e.g. camera parameters in case of image observations. The functional and the object model always need to be considered together and should be tuned accordingly, e.g. if the object model is rather course (like a cuboid representing a building) it is not necessary to use a very sophisticated set of camera parameters, on the other hand, if we demand very accurate results or a very high level of detail on the object, the functional model should be suitable to achieve these goals. $\boldsymbol{x}_{Object}$ and $\boldsymbol{x}_{Instrument}$ are combined in the vector $\boldsymbol{x}$ in the following.

The *stochastic model* describes the statistic properties of the observations. We know from our experience that observations are subject to some sort of error. For example, if the distance between two points is measured several times by different persons using different devices or methods, all observed distance values will be different. They will scatter around the 'true' value of the distance. Although we do not know this true value it is evident that the observations must be erroneous due to the scattering within the observations. Three types of errors are to be considered in this context: *random errors*, *gross errors* and *systematic errors*.

D.2.1 Error properties of observations

Random errors result from the accidental nature of the observations. The scattering in repeated observations is caused by the limitations of the measuring instrument and the user, the varying environmental influences where the observations take place, and the sharpness or definition uncertainty of the 3D-object itself. Therefore the result of a measurement process has to be considered as a *random variable*, with the actual observation being a *realization* or *sample* of it. Associated with a (continuous) random variable is a probability density function and thus a certain distribution. Usually this distribution of the measurement errors is assumed to be a *Gaußian* or *normal distribution*. However, the normal distribution is not generally required for errors to have 'random' behavior. Still, the normal distribution is a good choice because the repeated observations usually display a normal frequency distribution. It is also statistically well justified to assign the observations a normal distribution due to the *central limit theorem*, which states that data which are influenced by many small and unrelated random effects are approximately normally distributed.

The amount of the random errors or noise in a single observation can be quantified by the *variance* σ^2 or the *standard deviation* σ. The noise of and the mutual dependencies between a group of observations $\boldsymbol{m}$ can be quantified by the *covariance matrix* Σ_{mm}, which is a symmetric matrix. If two observations are stochastically independent then their covariance is zero. The inverse is not true in general, however, if the observations are normally distributed, then they are independent if their covariance is zero. If the covariance σ_{xy} between two observations is normalized by the respective standard deviations σ_x and σ_y, a measure called *correlation coefficient* is obtained. If all observations are uncorrelated then Σ_{mm} turns into a diagonal matrix. Data can be correlated in two ways: *mathematical correlation* (if the data results functionally from

other quantities; e.g. GPS-positions which are obtained from distance measurements to several satellites) and *physical correlation* (if certain environmental conditions like temperature or vapor fundamentally effect the data).

Gross errors, although occurring randomly, do not have any probability distribution and must be eliminated from the observations before computing the final results. When dealing with images they are caused by the carelessness of the user (who identifies a point wrongly from one image to another) or by the limitations of automatic (matching) algorithms. Usually gross errors or *blunders* are associated with a rather high magnitude. They, however, can also have rather small magnitude (in the order of the random errors) and then their detection turns out to be rather difficult.

Systematic errors actually are not a property of the observations. They result from errors in the functional model; i.e. the relations between the observations and the parameters are not correct. For example the disregard of existing non-linear distortion in the camera model will result in such systematic errors. The model errors must be detected before computing the final results and eliminated by a refinement of the functional model using additional parameters. The detection of small systematic errors can be done easier compared to small gross errors, since the model errors affect all data in the same way, thus resulting in a suspicious, i.e. systematic, pattern of errors, whereas small gross errors occur only individually.

D.3 Methods for parameter estimation

As we saw, due to the inevitable errors in the observations $\boldsymbol{m}$ we have to make more observations than necessary to minimize the influence of the errors on the resulting parameters $\boldsymbol{x}$ and the question arises, how to determine the unknowns in a *best* way given this redundant erroneous information.

Before we can do this, we have to define what we mean with 'best way'. Obviously some kind of general error or quality criterion has to be chosen in beforehand. A parameter estimator that either minimizes or maximizes this criterion can be regarded as a best way to determine the unknowns with respect to this criterion. The following three estimators maximizing/minimizing certain criteria seem appropriate:

- **Maximum Likelihood Estimator** The observations $\boldsymbol{m}$ are associated with a probability density function depending on the unknown parameters $\boldsymbol{x}$. The unknown parameters are determined in the way that the Likelihood function $L(\boldsymbol{m}|\boldsymbol{x})$ is maximized; i.e. the unknowns are determined such that the probability for the occurrence of the given sample of observations $\boldsymbol{m}$ is a maximum.

- **Best Unbiased Estimator** The unknown parameters $\boldsymbol{x}$ are determined in the way that the expectation of their estimation $E(\widehat{\boldsymbol{x}})$ and the expectation of their true value $E(\overset{*}{\boldsymbol{x}})$ are identical ($\rightarrow$ *unbiased* estimation) and that the expectation of the sum of the squared errors (i.e. the variance) is minimal ($\rightarrow$ *best* estimation). The expectations need to be considered because of the accidental nature of the errors in $\boldsymbol{m}$.

- **Least Squares Estimator** The unknown parameters $\boldsymbol{x}$ are determined in the way that the squared sum of the observations' residuals $\boldsymbol{v} = \boldsymbol{m} - s(\widehat{\boldsymbol{x}})$ is minimized: $\boldsymbol{v}^\top \boldsymbol{P}\boldsymbol{v} \rightarrow min$. With $s(\widehat{\boldsymbol{x}}) = \widehat{\boldsymbol{m}}$ being the estimated observations fitting to the estimated unknowns $\widehat{\boldsymbol{x}}$ and

P being a weight matrix for the observations $\boldsymbol{m}$ providing a sum of weighted squared residuals (see following text).

In contrast to the Best Unbiased and the Least Squares Estimator, the Maximum Likelihood Estimator requires a given probability density function for the observations.

For the least squares estimator a *weight matrix* P was introduced for the following reason: Due to the redundant erroneous information certain contradictions in the mathematical model will occur. For the least squares estimation these contradictions will be eliminated by the unknown parameters $\boldsymbol{x}$ and by altering (or improving) the given observations $\boldsymbol{m}$ by some residuals $\boldsymbol{v}$. *Weights* are used to reflect to what extent observations may be altered. Very accurate observations m_i (i.e. with small standard deviation σ_i) should be improved much less than observations with lower accuracy. Thus, for stochastically independent observations m_i, the weights are defined as $p_i = \sigma_0^2/\sigma_i^2$, with σ_0^2 being the *reference variance*. Its root is the *standard error of unit weight* (since an observation with $\sigma_i = \sigma_0$ has $w_i = 1$). All observations' weights are collected in the weight matrix P, which is a diagonal matrix for uncorrelated observations.

For stochastically dependent (thus correlated) observations individual weights can not be assigned, due to mutual dependencies. The weight matrix P is then defined as

$$P = Q_{mm}^{-1} = \sigma_0^2 \Sigma_{mm}^{-1} \tag{D.1}$$

Thus P is the inverse of the cofactor matrix Q_{mm} which is proportional to the observations' covariance matrix Σ_{mm}. If the observations are represented as homogenous entities, the covariance matrix Σ_{mm} will be singular; cf. [Heuel 2002]. Therefore, we will represent all observations inhomogenously and work with regular covariance matrices.

The functional relations $\boldsymbol{F}(\boldsymbol{m}, \boldsymbol{x}) = \boldsymbol{0}$ between observations $\boldsymbol{m}$ and (un)known parameters $\boldsymbol{x}$ are made up of two different kinds of relations: *condition equations* and *constraints*. Condition equations involve at least one observation and any number of parameters. Constraints do not include any observations and are entirely made up of parameters. Generally both of these relations can be made up of any (nonlinear) functions in $\boldsymbol{m}$ and $\boldsymbol{x}$ respectively. However, in general it is rather difficult to handle nonlinear estimation problems directly – in contrast to linear problems. Therefore a linearization of the original estimation problem is carried out by Taylor's series. This linearization needs an initial value $\boldsymbol{x}^0$ for the unknowns $\boldsymbol{x}$ which are iteratively refined by corrections $\Delta\boldsymbol{x}$.

If this linearization is done with the original functional relations $\boldsymbol{F}(\boldsymbol{m}, \boldsymbol{x}) = \boldsymbol{0}$ then one arrives at the so-called *Gauß-Markoff* (D.4) or *Gauß-Helmert* (D.5) models. The (iteratively refined) solution in these two models returns the Best Unbiased estimation for the unknowns (i.e. having minimal variance) and also minimizes the squared sum of the weighted observations' residuals $\boldsymbol{v}^\top P \boldsymbol{v}$. If the probability density function of the observations is known to be Gaußian, then the solution is also a Maximum Likelihood estimation.

Since the solution is independent of the choice of the criterion, the (weighted) Least Squares criterion will be considered in the following.

D.3.1 Parameterization

The object model and the functional model introduce parameters into our mathematical model. Most of these parameters will be unknown and a few will be known (i.e. constant). In a correct

manner only a few numbers can be considered as constant (e.g. the numbers e and π), because even internal parameters of a camera determined in a laboratory are only known with finite (admittedly very high) accuracy. So, in a rigorous mathematical model also the known parameters would be treated as stochastic variables. Consequently their covariance matrices would be considered also in the stochastic model.

It may be possible that a model for our (physical or geometrical) phenomenon can be represented by different sets of parameters. Among all possible sets of parameters, the so-called *minimal parameterizations* are distinguished, as they use the least possible number, referred to as u_0. For example, a 2D line can be represented in the following ways with the number of parameters given in brackets; adopted from [Heuel 2002]:

$$\begin{aligned}
\text{slope-intercept form (2)}: &\quad y = kx + m \\
\text{axis-intercept form (2)}: &\quad x/a + y/b = 1 \\
\text{angle-distance form (2)}: &\quad \cos\phi\, x + \cos\phi\, y + d = 0 \\
\text{homogenous form (3)}: &\quad ax + by + c = 0 \\
\text{point-angle form (3)}: &\quad \cos\phi\,(x - x_0) + \cos\phi\,(x - y_0) = 0 \\
\text{determinant form (4)}: &\quad \begin{vmatrix} x_1 & x_2 & x \\ y_1 & y_2 & y \\ 1 & 1 & 1 \end{vmatrix} = 0
\end{aligned}$$

Some of these representations use the minimal number of $u_0 = 2$ parameters, others are over-parameterized and use up to 4 parameters; however, some of the minimal representations can not represent all possible lines, e.g. vertical lines in the slope-intercept form. And therefore more than one map from all possible lines to a minimal parameterization may be needed; i.e. the two maps $y = kx + m$ and $x = k'y + m'$ represent all possible lines.

All of these representations return a valid 2D line for any choice of the parameters. A parameterization with this property is also termed *consistent*. In principle, a 2D line can also be represented as $y = fx^2 + gx + h$. This is a non-minimal representation, since 3 parameters are used, and it is also non-consistent, as only in the case $f = 0$ a valid 2D line is obtained.

If a non-minimal representation with u parameters is used, $u - u_0$ constraints must be satisfied by this parameterization. The purpose of these constraints in case of an already consistent parameterization is to prevent the singularity of the subsequent system of equations; e.g. for the homogenous form ax+by+c=0 the scale of the vector $(a, b, c)^\top$ must be fixed. In the case of an non-consistent representation, the constraints are required to make the parameterization consistent; e.g. $f = 0$ is the constraint required for $y = fx^2 + gx + h$ to represent a valid 2D line.

Note, that in the latter case the subsequent system of equations will be regular even without considering the constraints, but the result will not be valid as the additional parameters, i.e. the quadratic component f, will absorb some effects of the errors in the observations. However, in case of observations having only small errors, the observed points are practically collinear and therefore the quadratic component f is not determinable and the system will be nearly-singular.

If c constraints are considered in the parameterization with u parameters, we will refer to the number $u - c$ as the *number of essential parameters*[1] u_e. Only if $u_e = u_0$, the parameterization together with the constraints is consistent. For example, if the trifocal tensor is parameterized by its 27 elements and only the scaling constraint is considered, then $u_e = 26$ and $u_0 = 18$.

[1]The term *degrees of freedom* might be a better choice for the number u_e, however, this term is already defined

D.4 The Gauß-Markoff model

The *Gauss-Markoff model* model is defined as $\widehat{\boldsymbol{m}} = \boldsymbol{f}(\widehat{\boldsymbol{x}})$; i.e. the original nonlinear relations $\boldsymbol{F}(\widehat{\boldsymbol{m}}, \widehat{\boldsymbol{x}}) = \boldsymbol{0}$ only include condition equations where each observation is expressed by some set of parameters and therefore $m = c > u$). The Taylor linearization in the Gauß-Markoff model is given by:

$$\begin{aligned} \widehat{\boldsymbol{m}} = \boldsymbol{m} + \boldsymbol{v} = \boldsymbol{f}(\boldsymbol{x}^0 + \Delta\boldsymbol{x}) \rightarrow \boldsymbol{v} &= \boldsymbol{A}\Delta\boldsymbol{x} - \boldsymbol{w} \\ \boldsymbol{w} &= \boldsymbol{m} - \boldsymbol{f}(\boldsymbol{x}^0) \end{aligned} \tag{D.2}$$

The term 'Gauss-Markoff model' is used, because [Gauß 1809] applied it together with the Likelihood method to derive the Least Squares method and showed that this way a Best Estimation is obtained. Later [Markoff 1912] used this model to determine the unknowns using the Best Unbiased Estimator. This model is also called *adjustment by indirect observations* and plays the mayor role in technical disciplines since in many situations a single (physical) observation returned by some instrument can be related easily to a set of unknown physical or geometrical parameters. So e.g. this model is used for *bundle adjustment*, where the measurements in an image can be expressed easily by the orientation elements of the image (e.g. the projection matrix) and the corresponding 3D-points.

The sum of the weighted squared residuals Ω is given by

$$\Omega = \boldsymbol{v}^\top \boldsymbol{P}\boldsymbol{v} \rightarrow min. \tag{D.3}$$

By replacing $\boldsymbol{v}$ in (D.3) with (D.2) and setting the partial derivative $\partial\Omega/\partial\Delta\boldsymbol{x} = \boldsymbol{0}$ the correction $\Delta\boldsymbol{x}$ for the unknown parameters are determined by:

$$\Delta\boldsymbol{x} = (\boldsymbol{A}^\top \boldsymbol{P}\boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{P}\boldsymbol{w} = \boldsymbol{N}^{-1}\boldsymbol{A}^\top \boldsymbol{P}\boldsymbol{w} \tag{D.4}$$

The system (D.4) is also termed *system of normal equations*. This is due to the fact that the residual vector $\boldsymbol{v}$ is orthogonal to the column space of the Jacobian $\boldsymbol{A}$: $\boldsymbol{A}^\top \boldsymbol{v} = \boldsymbol{0}$. If the model is over-parameterized the Jacobian $\boldsymbol{A}$ will be deficient in rank and (D.4) needs to be solved using the pseudo-inverse $\boldsymbol{N}^+$.

The unknown σ_0 can be estimated by

$$\widehat{\sigma}_0 = \sqrt{\frac{\boldsymbol{v}^\top \boldsymbol{P}\boldsymbol{v}}{m-u}} \tag{D.5}$$

The covariance matrix $\Sigma_{\widehat{x}\widehat{x}}$ of the estimated unknown parameters $\widehat{x}$ is computed by

$$\Sigma_{\widehat{x}\widehat{x}} = \widehat{\sigma}_0^2 Q_{\widehat{x}\widehat{x}} \tag{D.6}$$

$$Q_{\widehat{x}\widehat{x}} = (\boldsymbol{A}^\top \boldsymbol{P}\boldsymbol{A})^{-1} = \boldsymbol{N}^{-1} \tag{D.7}$$

The covariance matrices of the other quantities can be found in [Mikhail 1976] or [Koch 1999].

In [Gill and Murray 1978] the solution (D.4) is referred to as *Gauss-Newton's method*, which differs from another one called *Newton's method*. The difference is caused by the argument

in adjustment theory, where it used for the number of redundant observations $m + c - u$, given m observations for u unknown parameters and c constraints. To avoid confusion with the notion of degrees of freedom used in the geometric part of this thesis, we will not use this term in this chapter on adjustment at all.

of the Taylor approximation, which is $f(x)$ for Gauß-Newton's and Ω for Newton's method. Therefore the latter takes the Hessian matrices of $f(x)$ into account. Due to [Gill and Murray 1978] the Gauß-Newton method converges to the real value $\overset{*}{x}$ as long as the Jacobian A is regular and $|m - f(x)|$ is small.

Another modification of this model is the the so-called *Levenberg-Marquardt algorithm* which uses *augmented normal equations* by replacing N in (D.4) with N' where $N'_{ii} = (1 + \lambda)N_{ii}$ for some small λ. Due to [Hartley and Zisserman 2001] the Levenberg-Marquardt algorithm moves seamlessly between the Gauß-Newton method, which will cause rapid convergence in the vacinity of the solution, and a descent approach, which will guarantee a decrease in the cost function Ω when the going is difficult.

D.5 The Gauß-Helmert model

The *Gauss-Helmert model* is defined as a functional model $F(\widehat{m}, \widehat{x}) = 0$, which is made up of the condition equations $f(\widehat{m}, \widehat{x}) = 0$ and also may include additional constraints $g(\widehat{x}, \widehat{y}) = 0$, with y being some sort of additional unknown parameters, which may be needed to set up the constraints, but do not take part in the condition equations. Since it makes no restrictions on the functional model the Gauss-Helmert model is also called *general case of least squares adjustment* [Helmert 1872]. The Taylor linearization of the Gauß-Helmert model is given by:

$$\begin{aligned} \text{conditions: } f(\widehat{m}, \widehat{x}) = 0 \quad &\rightarrow \quad A\Delta x + Bv - w = 0 \\ w &= -f(m, x^0) - B(m - m^0) \\ \text{constraints: } g(\widehat{x}, \widehat{y}) = 0 \quad &\rightarrow \quad C\Delta x + D\Delta y - t = 0 \\ t &= -g(x^0, y^0) \end{aligned} \tag{D.8}$$

Observe, that the vector w in (D.8) depends on the original observations m *and* the iteratively refined approximations m^0 for their estimates $\widehat{m}$.

With m independent observations m, u ($\geq u_0$) unknown parameters x, c_1 independent conditions, c_2 independent constraints set up with s additional independent unknowns y the Jacobians resp. vectors in (D.8) are of the following order:

x: $u \times 1$	v: $m \times 1$	y: $s \times 1$
A: $c_1 \times u$, rank $= u_0$	B: $c_1 \times m$, rank $= c_1$	w: $c_1 \times 1$
C: $c_2 \times u$, rank $= c_2 - s$	D: $c_2 \times s$, rank $= s$	t: $c_2 \times 1$

And it holds:

$$\begin{aligned} m &\geq c_1 \\ c_1 + c_2 &\geq u + s \\ c_2 &\leq u + s \\ c_2 &\geq s \\ u_e &= u + s - c_2 \end{aligned}$$

Using Lagrange multipliers $\boldsymbol{\lambda}$ (a c_1-dimensional vector) and $\boldsymbol{\mu}$ (a c_2-dimensional vector) the criterion Ω to be minimized is given by

$$\begin{aligned}\Omega &= \boldsymbol{v}^\top \boldsymbol{P}\boldsymbol{v} \\ &\quad + 2\boldsymbol{\lambda}^\top(\boldsymbol{A}\Delta\boldsymbol{x} + \boldsymbol{B}\boldsymbol{v} - \boldsymbol{w}) \\ &\quad + 2\boldsymbol{\mu}^\top(\boldsymbol{C}\Delta\boldsymbol{x} + \boldsymbol{D}\Delta\boldsymbol{y} - \boldsymbol{t}) \rightarrow min. \end{aligned} \tag{D.9}$$

The partial derivatives of Ω are set to $\mathbf{0}$:

$$\begin{aligned}\partial\Omega/\partial\boldsymbol{v} &= 2\boldsymbol{P}\boldsymbol{v} + 2\boldsymbol{B}^\top\boldsymbol{\lambda} = \mathbf{0} \\ &\quad \boldsymbol{v} = -\boldsymbol{P}^{-1}\boldsymbol{B}^\top\boldsymbol{\lambda} \end{aligned} \tag{D.10}$$

$$\begin{aligned}\partial\Omega/\partial\Delta\boldsymbol{x} &= 2\boldsymbol{A}^\top\boldsymbol{\lambda} + 2\boldsymbol{C}^\top\boldsymbol{\mu} = \mathbf{0} \\ \partial\Omega/\partial\Delta\boldsymbol{y} &= 2\boldsymbol{D}^\top\boldsymbol{\mu} = \mathbf{0} \\ \partial\Omega/\partial\boldsymbol{\lambda} &= \boldsymbol{A}\Delta\boldsymbol{x} + \boldsymbol{B}\boldsymbol{v} - \boldsymbol{w} = \mathbf{0} \\ &\quad \boldsymbol{\lambda} = (\boldsymbol{B}\boldsymbol{P}^{-1}\boldsymbol{B}^\top)^{-1}(\boldsymbol{A}\Delta\boldsymbol{x} - \boldsymbol{w}) \\ &\quad \boldsymbol{\lambda} = \boldsymbol{W}(\boldsymbol{A}\Delta\boldsymbol{x} - \boldsymbol{w}) \end{aligned} \tag{D.11}$$

$$\partial\Omega/\partial\boldsymbol{\mu} = \boldsymbol{C}\Delta\boldsymbol{x} + \boldsymbol{D}\Delta\boldsymbol{y} - \boldsymbol{t} = \mathbf{0}$$

The matrix $\boldsymbol{W} = (\boldsymbol{B}\boldsymbol{P}^{-1}\boldsymbol{B}^\top)^{-1}$ in (D.11), obtained by error propagation from $\boldsymbol{P}$, is the weight matrix of the conditions. This matrix is regular if $\boldsymbol{P}$ is regular (i.e. inhomogenous observations are used) and if $\boldsymbol{B}$ has full row rank (i.e. the conditions $\boldsymbol{f}(\boldsymbol{m}, \boldsymbol{x})$ are linearly independent). If the unknown vector of Lagrange multipliers $\boldsymbol{\lambda}$ is eliminated with (D.11) then the following system of normal equations is obtained:

$$\begin{bmatrix} \boldsymbol{A}^\top\boldsymbol{W}\boldsymbol{A} & \mathbf{0} & \boldsymbol{C}^\top \\ \mathbf{0} & \mathbf{0} & \boldsymbol{D}^\top \\ \boldsymbol{C} & \boldsymbol{D} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{y} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}^\top\boldsymbol{W}\boldsymbol{w} \\ \mathbf{0} \\ \boldsymbol{t} \end{pmatrix} \tag{D.12}$$

If the conditions and constraints are properly set up the system (D.12) will be regular. If $\boldsymbol{B}$ has not full row rank (i.e. the conditions are dependent) then $\boldsymbol{W}$ is obtained using the pseudo-inverse of $\boldsymbol{B}\boldsymbol{P}^{-1}\boldsymbol{B}^\top$. The pseudo-inverse is briefly summarized in section B.2 on page 178. Further if the model is overparameterized with a consistent set of parameters and the dependencies are not (entirely) considered in the constraints, then the whole system (D.12) needs to be solved using its pseudo-inverse.

The Lagrange multipliers $\boldsymbol{\lambda}$ and further the residuals $\boldsymbol{v}$ can be computed from the corrections $\Delta\boldsymbol{x}$ using (D.11) and (D.10) respectively. This way better approximations for the next iteration step are obtained. Since the Jacobians $\boldsymbol{A}$ and $\boldsymbol{B}$ depend on the parameters *and* the observations, $\boldsymbol{A}$ and $\boldsymbol{B}$ for the iteration (i+1) should be evaluated at $(\boldsymbol{x}^0_{i+1}, \boldsymbol{m}^0_{i+1}) = (\boldsymbol{x}^0_i + \Delta\boldsymbol{x}_i, \boldsymbol{m} + \boldsymbol{v}_i)$. However, in practise the iteration on the residuals may be skipped. This can be admissible if the errors in the observations are not too large; i.e. the given observations $\boldsymbol{m}$ are sufficiently close to the final estimates. If the residual iteration is neglected (and if no constraints are in the model), then this method is identical to the minimization of the *Sampson approximation to reprojection error* presented in [Hartley and Zisserman 2001]. However, one must be aware of the fact, that if the iteration on the residuals is skipped, the final check of the conditions will not be fulfilled; i.e. $\boldsymbol{f}(\widehat{\boldsymbol{m}}, \widehat{\boldsymbol{x}}) \neq \mathbf{0}$.

The unknown σ_0 can be estimated by

$$\widehat{\sigma}_0 = \sqrt{\frac{\boldsymbol{v}^\top \boldsymbol{P} \boldsymbol{v}}{c_1 + c_2 - u - s}} \tag{D.13}$$

The system (D.12) can also be represented by:

$$\begin{pmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{y} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q}_{11} & \boldsymbol{Q}_{12} & \boldsymbol{Q}_{13} \\ \boldsymbol{Q}_{12}^\top & \boldsymbol{Q}_{22} & \boldsymbol{Q}_{23} \\ \boldsymbol{Q}_{13}^\top & \boldsymbol{Q}_{23}^\top & \boldsymbol{Q}_{33} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}^\top \boldsymbol{W}^{-1} \boldsymbol{w} \\ \boldsymbol{0} \\ \boldsymbol{t} \end{pmatrix} \tag{D.14}$$

Using (D.14) the covariance matrix $\Sigma_{\widehat{x}\widehat{x}}$ of the estimated unknown parameters $\widehat{x}$, with $\operatorname{rank}(\Sigma_{xx}) = u_0$, is computed by

$$\Sigma_{\widehat{x}\widehat{x}} = \widehat{\sigma}_0^2 \boldsymbol{Q}_{11}. \tag{D.15}$$

Note: By considering or neglecting the matrices $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$, the Gauß-Helmert model allows the unknowns to be determined by minimizing measurement error (with $\boldsymbol{B}$) resp. algebraic error (without $\boldsymbol{B}$), in a constrained way (with $\boldsymbol{C}$ and $\boldsymbol{D}$) or unconstrained way (without $\boldsymbol{C}$ and $\boldsymbol{D}$).

Appendix E

MATLAB File Listings

Two proofs in this thesis were done using MATLAB V.5 – a programming language for technical computing. The respective listings of the applied script files (the so-called *m-files*) are given here.

E.1 Vanishing derivatives for the essential matrix constraint of identical singular values

Following proposition 6.5 on page 49 a matrix E_{12} is an essential matrix if and only if its singular values $\{\sigma_1, \sigma_2, \sigma_3\}$ satisfy the following condition

$$\sigma_1 = \sigma_2 \quad \text{and} \quad \sigma_3 = 0.$$

The first condition $\sigma_1 = \sigma_2$ expressed in the elements of E_{12} reads as

$$(\mathbf{e}_1^\top \mathbf{e}_1 + \mathbf{e}_2{}^\top \mathbf{e}_2 + \mathbf{e}_3^\top \mathbf{e}_3)^2 = 4(\mathbf{e}_1^{\star\top} \mathbf{e}_1^\star + \mathbf{e}_2^{\star\top} \mathbf{e}_2^\star + \mathbf{e}_3^{\star\top} \mathbf{e}_3^\star) \tag{E.1}$$

with $\mathsf{E}_{12}^\star = [\mathbf{e}_1^\star, \mathbf{e}_2^\star, \mathbf{e}_3^\star] = [\mathbf{e}_2 \times \mathbf{e}_3, \mathbf{e}_3 \times \mathbf{e}_1, \mathbf{e}_1 \times \mathbf{e}_2]$ being the co-factor matrix of E_{12}.

With the following m-file it is shown, that the differentials of the equation E.1 with respect to the 9 elements of E_{12} vanish if E_{12} is a valid essential matrix. In practical applications this means, that starting from approximate values for E_{12} the iteration sequence crashes in the moment of enditeration.

```
% This MATLAB file shows the following:
% Let C represent the constraint that the essential matrix (E) has two identical singular
% values.
% Then the differntials of C with respect to the 9 elements of E vanish if the 9 elements
% make up a valid essential matrix; i.e. when the adjustment in the Gauss-Helmert model
% is enditerated.

clear;                                          % delete old variables
syms e11 e12 e13 e21 e22 e23 e31 e32 e33 real;  % 9 elements of E defined as real symbols
e1=[e11;e12;e13];                               % the columns of E
e2=[e21;e22;e23];
e3=[e31;e32;e33];

E=[e1 e2 e3];                                   % E as a general 3x3 matrix
```

```
                                                % constraint C=0
leftside=(e1'*e1+e2'*e2+e3'*e3)^2;
rightside=4*([cross(e1,e2)]'*cross(e1,e2)+
             [cross(e1,e3)]'*cross(e1,e3)+
             [cross(e2,e3)]'*cross(e2,e3));
C=leftside-rightside;
                                                % derivatives of C with respect to the
                                                % elements of E
deriv_C11=   diff(C,'e11');
deriv_C12=   diff(C,'e12');
deriv_C13=   diff(C,'e13');
deriv_C21=   diff(C,'e21');
deriv_C22=   diff(C,'e22');
deriv_C23=   diff(C,'e23');
deriv_C31=   diff(C,'e31');
deriv_C32=   diff(C,'e32');
deriv_C33=   diff(C,'e33');



                                        % a valid essential matrix  with 5 DOF
syms b1 b2 b3 real;                     % elements of baseline
B=[ 0 -b3 b2; b3 0 -b1; -b2 b1 0];      % axiator of baseline
syms s1 s2 s3 real;                     % parameters for spatial rotation
S=[ 0 -s3 s2; s3 0 -s1; -s2 s1 0];
R=inv(eye(3)-S)*(eye(3)-S);             % rotation matrix

E5dof=B*R;                              % the valid essential matrix


                                        % assign the 9 valid elements to the
                                        % general 9 elements
e11=E5dof(1,1);e12=E5dof(1,2);e13=E5dof(1,3);
e21=E5dof(2,1);e22=E5dof(2,2);e23=E5dof(2,3);
e31=E5dof(3,1);e32=E5dof(3,2);e33=E5dof(3,3);


                                       % substitute the valid elements in the deriv.
subst_deriv_C11=subs(deriv_C11);
subst_deriv_C12=subs(deriv_C12);
subst_deriv_C13=subs(deriv_C13);
subst_deriv_C21=subs(deriv_C21);
subst_deriv_C22=subs(deriv_C22);
subst_deriv_C23=subs(deriv_C23);
subst_deriv_C31=subs(deriv_C31);
subst_deriv_C32=subs(deriv_C32);
subst_deriv_C33=subs(deriv_C33);

simpl_deriv_C=zeros(1,9);                 % simplify the expressions of the deriv.
simpl_deriv_C(1)=simplify(subst_deriv_C11);
simpl_deriv_C(2)=simplify(subst_deriv_C12);
simpl_deriv_C(3)=simplify(subst_deriv_C13);
simpl_deriv_C(4)=simplify(subst_deriv_C21);
simpl_deriv_C(5)=simplify(subst_deriv_C22);
simpl_deriv_C(6)=simplify(subst_deriv_C23);
simpl_deriv_C(7)=simplify(subst_deriv_C31);
simpl_deriv_C(8)=simplify(subst_deriv_C32);
simpl_deriv_C(9)=simplify(subst_deriv_C33);
simpl_deriv_C                            % output of simplified derivatives,
                                         % which turn out to be all zero
```

E.2 Independency of the constraints derived from the homography slices

In section 7.6.3 on page 103 we found the following constraints from the J_j homographies: The three rank constraints (7.52) on page 100 and the single epipolar constraint (7.53) on page 101, these four constraints were also derived from the I_i correlations. Further we found the collinearity constraints (7.58) on page 105 and the concurrency constraint (7.59) on page 105. These eight constraints must be independent to represent a minimal set of constraints.

If the J_j homographies satisfy the first four constraints, then the parameterization 7.56 on page 102 of the I_i correlations can be applied, which represents the J_j homographies as

$$\begin{aligned}
\mathsf{J}_1 &= \left[\mathbf{s}_1''', \mathbf{s}_2''', \mathbf{s}_3'''\right], \\
\mathsf{J}_2 &= \left[v_1\mathbf{s}_1''' + m_1\mathbf{e}_1''', v_2\mathbf{s}_2''' + m_2\mathbf{e}_1''', v_3\mathbf{s}_3''' + m_3\mathbf{e}_1'''\right], \\
\mathsf{J}_3 &= \left[w_1\mathbf{s}_1''' + n_1\mathbf{e}_1''', w_2\mathbf{s}_2''' + n_2\mathbf{e}_1''', w_3\mathbf{s}_3''' + n_3\mathbf{e}_1'''\right].
\end{aligned}$$

Using this parameterization the other four constraints, i.e. the collinearity constraints (7.58) and the concurrency constraint (7.59), can be set up. If they are satisfied, all eight constraints are dependent, if they are not satisfied, they are independent.

Note, that there is a some uncertainty in this proof, as the desired goal of independent eight constraints would also result if the collinearity constraints (7.58) and the concurrency constraint (7.59) were incorrectly implemented. Therefore the implementation is also checked with the consistent parameterization (7.57), which is obtained by $v_1 = v_2 = v_3 = v$ and $w_1 = w_2 = w_3 = w$. For this consistent parameterization the collinearity constraints (7.58) and the concurrency constraint (7.59) must be satisfied.

```
% ++++ Minimal set of constraints via the J-matrices ++++
% The J-matrices are parameterized so that they satisfy the constraints in image 3.
% With this parameterization, the constraints in image 1 are checked.
% It should result that the constraints in image 1 are not satisfied
% and therefore all 8 constraints are independent

clear

                                          % parameterization of the J-matrices so that
                                          % they satisfy the constraints in image 3
syms s11 s12 s13 s21 s22 s23 s31 s32 s33;
syms e31x e31y e31z;                      % epipole in image 3 with respect to image 1

s1=[s11;s12;s13];
s2=[s21;s22;s23];
s3=[s31;s32;s33];

e31=[e31x;e31y;e31z];
syms v1 v2 v3 w1 w2 w3;
syms m1 m2 m3 n1 n2 n3;

J1=[s1,s2,s3];                             % parameterization of J-matrices
J2=[v1*s1+m1*e31,v2*s2+m2*e31,v3*s3+m3*e31];
J3=[w1*s1+n1*e31,w2*s2+n2*e31,w3*s3+n3*e31];


                                           % constraints in image 1
col12=det([cross(J1(1,:),J1(2,:));
           cross(J2(1,:),J2(2,:));
           cross(J3(1,:),J3(2,:))]);       % collinearity on rho_33
```

```
col13=det([cross(J1(1,:),J1(3,:));
           cross(J2(1,:),J2(3,:));
           cross(J3(1,:),J3(3,:))]);          % collinearity on rho_32
col23=det([cross(J1(3,:),J1(2,:));
           cross(J2(3,:),J2(2,:));
           cross(J3(3,:),J3(2,:))]);          % collinearity on rho_31
                                              % concurrency of rho_33, rho_32, rho_31
conc=det([cross(cross(J1(1,:),J1(2,:)),cross(J2(1,:),J2(2,:)))
          cross(cross(J1(1,:),J1(3,:)),cross(J2(1,:),J2(3,:)))
          cross(cross(J1(3,:),J1(2,:)),cross(J2(3,:),J2(2,:)))]);

[R,HOW] =simple(col12)                        % output is different from zero --> independency
[R,HOW] =simple(col13)
[R,HOW] =simple(col23)
[R,HOW] =simple(conc)

% check: constraints in image 1 should be satisfied if consistent parameterization is used
syms v w;
v1=v;v2=v;v3=v;
w1=w;w2=w;w3=w;

J1=[s1,s2,s3];                                % parameterization of valid J-matrices
J2=[v1*s1+m1*e31,v2*s2+m2*e31,v3*s3+m3*e31];
J3=[w1*s1+n1*e31,w2*s2+n2*e31,w3*s3+n3*e31];

                                              % constraints in image 1
col12=det([cross(J1(1,:),J1(2,:));
           cross(J2(1,:),J2(2,:));
           cross(J3(1,:),J3(2,:))]);
col13=det([cross(J1(1,:),J1(3,:));
           cross(J2(1,:),J2(3,:));
           cross(J3(1,:),J3(3,:))]);
col23=det([cross(J1(3,:),J1(2,:));
           cross(J2(3,:),J2(2,:));
           cross(J3(3,:),J3(2,:))]);
conc=det([cross(cross(J1(1,:),J1(2,:)),cross(J2(1,:),J2(2,:)))
          cross(cross(J1(1,:),J1(3,:)),cross(J2(1,:),J2(3,:)))
          cross(cross(J1(3,:),J1(2,:)),cross(J2(3,:),J2(2,:)))]);

[R,HOW] =simple(col12)                        % output is  zero --> check of implementation
[R,HOW] =simple(col13)
[R,HOW] =simple(col23)
[R,HOW] =simple(conc)
```

References

Avidan, S. and A. Shashua (1996). Tensorial transfer: On the representation of $N > 3$ views of a 3D scene. In *Proc. of the ARPA Image Understanding Workshop, Palm Springs, Feb. 1996*.

Avidan, S. and A. Shashua (1997). Novel view synthesis in tensor space. In *Conference on Computer Vision and Pattern Recognition (CVPR '97), June 17-19, 1997, San Juan, Puerto Rico*, pp. 1034–1040. IEEE Computer Society.

Brand, L. (1966). *Vector and Tensor Analysis, 10. Ed.* John Wiley & Sons, Inc.

Bronstein, I. and K. Semendjajew (1991). *Taschenbuch der Mathematik, 25. Auflage*. B.G. Teubner Verlagsgesellschaft, Stuttgart, Leipzig und Verlag Nauka, Moskau.

Buchanan, T. (1992). Critical sets for 3d reconstruction using lines. In *European Conference on Computer Vision, Proceedings*, Lecture Notes in Computer Science, pp. 730–738. Springer.

Canterakis, N. (2000). A minimal set of constraints for the trifocal tensor. In D. Vernon (Ed.), *Computer Vision - ECCV 2000, 6th European Conference on Computer Vision, Dublin, Ireland, June 26 - July 1, 2000, Proceedings, Part I*, Volume 1842 of *Lecture Notes in Computer Science*, pp. 84–99. Springer.

Carlsson, S. (1997). *Geometry and Algebra of Projective Views*. Lecture Notes, Nov. 1997, http://www.nada.kth.se/ stefanc/.

Cramer, M. (2001). Genauigkeitsuntersuchungen zur GPS/INS-Integration in der Aerophotogrammetrie. *Deutsche Geodätische Kommission München, Reihe C, Heft Nr. 537*.

Faugeras, O. (1992). Camera self-calibration: Theory and experiments. In *European Conference on Computer Vision, Proceedings*, Lecture Notes in Computer Science, pp. 321–334. Springer.

Faugeras, O. and Q. Luong (2001). *The Geometry of Multiple Images*. USA: MIT Press.

Faugeras, O. and S. Maybank (1990). Motion from point matches : multiplicity of solutions. *International Journal of Computer Vision 4*, 225–246.

Faugeras, O. and B. Mourrain (1995). On the geometry and algebra of the point and line correspondences between N images. In *Proceedings of ICCV95*, pp. 951–956.

Fischler, M. and R. Bolles (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM 24*(6), 381–395.

Fitzgibbon, A. (2001). Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Förstner, W. (2000). On weighting and choosing constraints for optimally reconstructing the geometry of image triplets. In D. Vernon (Ed.), *Computer Vision - ECCV 2000, 6th European*

Conference on Computer Vision, Dublin, Ireland, June 26 - July 1, 2000, Proceedings, Part II, Volume 1843 of *Lecture Notes in Computer Science*, pp. 669–684. Springer.

Förstner, W. (2003). private communications.

Gauß, C. (1809). *Theoria motus corporum coelestium*. Hamburg: Perthes und Besser.

Gill, P. and W. Murray (1978). Algorithms for the solution of the nonlinear least squares problem. *SIAM J. Numer. Anal. 15*, 977–992.

Golub, G. and C. VanLoan (1996). *Matrix computations 3. Ed.* Baltimore, Md.: Johns Hopkins Univ. Press.

Hartley, R. (1992). Estimation of relative camera positions for uncalibrated cameras. In *Computer Vision - ECCV 1992*, Lecture Notes in Computer Science, pp. 579–587. Springer.

Hartley, R. (1994a). Lines and points in three views: A unified approach. In *ARPA Image Understanding Workshop*, pp. 1009–1016.

Hartley, R. (1994b). Projective reconstruction from line correspondences. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pp. 903–907.

Hartley, R. (1995a). In defence of the 8-point algorithm. In *Proceedings of the 5th International Conference on Computer Vision*, pp. 1064–1070. IEEE Computer Society Press.

Hartley, R. (1995b). Multilinear relationships between coordinates of corresponding image points and lines. In *Proceedings of the International Workshop on Computer Vision and Applied Geometry, International Sophus Lie Center, Nordfjordeid, Norway, August 1995*. Vanderbilt University Press.

Hartley, R. (1997). Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision 22*, 125–140.

Hartley, R. (1998). Computation of the quadrifocal tensor. In *Computer Vision - ECCV 1998, 5th European Conference on Computer Vision, Proceedings*, Volume 1406 of *Lecture Notes in Computer Science*, pp. 20–35. Springer.

Hartley, R. (2000). Ambiguous configurations for 3-view projective reconstruction. In D. Vernon (Ed.), *Computer Vision - ECCV 2000, 6th European Conference on Computer Vision, Dublin, Ireland, June 26 - July 1, 2000, Proceedings, Part I*, Volume 1842 of *Lecture Notes in Computer Science*, pp. 922–935. Springer.

Hartley, R., R. Gupta, and T. Chang (1992). Stereo from uncalibrated cameras. In *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 761–764. Urbana Champaign, IL. IEEE.

Hartley, R. and A. Zisserman (2001). *Multiple View Geometry in Computer Vision, Reprinted Edition*. Cambridge, UK: Cambridge University Press.

Helmert, F. R. (1872). *Die Ausgleichungsrechnung nach der Methode der Kleinsten Quadrate*. Leipzig: Teubner.

Heuel, S. (2002). *Statistical Reasoning in Uncertain Projective Geometry for Polyhedral Object Reconstruction*. Ph. D. thesis, Bonn University.

Heyden, A. (2000). Tensorial properties of multilinear constraints. *Mathematical Methods in the Applied Sciences 23*, 169–202.

Huang, T. and O. Faugeras (1989). Some properties of the E-matrix in two view motion estimation. *IEEE Transaction of Pattern Analysis and Machine Intelligence 11*, 1310–1312.

Joshi, A. (1975). *Matrices and tensors in physics*. New Delhi: Wiley.

Klingbeil, E. (1966). *Tensorrechnung für Ingenieure*. Mannheim: Bibliogr. Inst.

Koch, K. R. (1999). *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer.

Kraus, K. (1993). *Photogrammetry, Volume 1: Fundamentals and Standard Processes, 4. Ed.* Dümmler/Bonn.

Kraus, K. (1997). *Photogrammetry, Volume 2: Advanced Methods and Applications, 4. Ed.* Dümmler/Bonn.

Lasenby, J. and A. Lasenby (1998). Estimating tensors for matching over multiple views. *Philosophical Transactions: Mathematical, Physical & Engineering Sciences 356*(1740), 1267–1282.

Longuet-Higgins, H. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature September*, 133–135.

Luong, Q. (1992). *Matrice fondamentale et calibration visuelle sur l'environnement: vers une plus grande autonomie des syst'emes robotiques*. Ph. D. thesis, University of Paris XI, Orsay, France.

Luong, Q. and O. Faugeras (1996). The fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision 17*, 43–76.

Markoff, A. (1912). *Wahrscheinlichkeitsrechnung*. Leibzig: Teubner.

Maybank, S. (1993). *Theory of Reconstruction from Image Motion*. Springer.

Mayer, H. (2002). Estimation of and view synthesis with the trifocal tensor. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 3A*, ISPRS-Comm. III Symposium, Graz, 9. - 13. September 2002, pp. 211–217.

Mikhail, E. (1962). Use of triplets in analytical aerotriangulation. *Photogr. Eng. 28*.

Mikhail, E. (1976). *Observations And Least Squares*. New York: IEP-A Dun-Donnelley.

Niini, I. (1994). Relative orientation of multiple images using projective singular correlation. Master's thesis, Helsinki University of Technology, Finland.

Niini, I. (2000). Photogrammetric block adjustment based on singular correlation. *Acta Polytechnica Scandinavica, Civil Engineering and Building Construction Series No. 120. Finnish Academies of Technology, Espoo*.

Papadopoulo, T. and O. Faugeras (1998). A new characterization of the trifocal tensor. In H. Burkhardt and B. Neumann (Eds.), *Computer Vision - ECCV 1998, 5th European Conference on Computer Vision, Proceedings*, Volume 1406 of *Lecture Notes in Computer Science*. Springer.

Philip, J. (1996). A non-iterative algorithm for determining all essential matrices corresponding to five point pairs. *Photogrammetric Record 15(88)*, 589–599.

Quan, L. (1994). Invariants of 6 points from 3 uncalibrated images. In J. Eklundh (Ed.), *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden, Volume II*, pp. 459–470. Springer.

Ressl, C. (2002). A minimal set of constraints and a minimal parameterization for the trifocal tensor. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 3A*, ISPRS-Comm. III Symposium, Graz, 9. - 13. September 2002, pp. 277–282.

Rinner, K. (1942). Studie über eine rechnerische Lösung für die gegenseitige Orientierung photogrammetrischer Aufnahmen. *Photogrammetria 5*, 41–54.

Rinner, K. (1956). Eine allgemeine analytische Lösung des Folgebildanschlusses. *Österreichische Zeitschrift für Vermessugswesen 1*, 4–9.

Rinner, K. (1963). Studien über eine allgemeine, voraussetzungslose Lösung des Folgebildanschlusses. *Sonderheft 23 der Österreichischen Zeitschrift für Vermessugswesen*.

Rinner, K. and R. Burkhardt (1972). Photogrammetrie. In W. Jordan, O. Eggert, and M. Kneissl (Eds.), *Handbuch der Vermessungskunde, 10. Ed., Band IIIa/2*. Stuttgart: Metzler'sche Verlagsbuchhandlung.

Schmidt, K. and G. Trenkler (1998). *Moderne Matrix-Algebra – mit Anwendungen in der Statistik*. Berlin: Springer.

Shashua, A. (1995a). Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence 17(8)*, 779–789.

Shashua, A. (1995b). Multiple-view geometry and photometry. In S. Li, D. Mital, E. Teoh, and H. Wang (Eds.), *Recent Developments in Computer Vision, Second Asian Conference on Computer Vision, ACCV '95, Singapore, December 5-8, 1995, Invited Session Papers*, Volume 1035 of *Lecture Notes in Computer Science*. Springer.

Shashua, A. and M. Werman (1995). Trilinearity of three perspective views and its associated tensor. In *Proceedings of the International Conference on Computer Vision, June 1995*, pp. 920–925.

Shashua, A. and L. Wolf (1994). Trilinearity in visual recognition by alignment. In J.-O. Eklundh (Ed.), *Computer Vision - ECCV'94, Third European Conference on Computer Vision, Stockholm, Sweden, May 2-6, 1994, Proceedings, Volume I*, Volume 800 of *Lecture Notes in Computer Science*, pp. 479–484. Springer.

Shashua, A. and L. Wolf (2000). On the structure and properties of the quadrifocal tensor. In D. Vernon (Ed.), *Computer Vision - ECCV 2000, 6th European Conference on Computer Vision, Dublin, Ireland, June 26 - July 1, 2000, Proceedings, Part I*, Volume 1842 of *Lecture Notes in Computer Science*. Springer.

Spetsakis, M. E. and Y. Aloimonos (1990). A unified theory of structure from motion. In *Proceedings of a Workshop held in Pittsburgh, Pennsylvania, Sept.11-13,1990*, pp. 271–283.

Spiegel, M. (1959). *Vector analysis and an introduction to tensor analysis*. New York: Schaum Publishing Co.

Stefanovic, P. (1973). Relative orientation - a new approach. *ITC-Journal 3*, 417–448.

Stein, G. and A. Shashua (1998). On degeneracy of linear reconstruction from three views: Linear line complex and applications. In *Computer Vision - ECCV 1998, 5th European Conference on Computer Vision, Proceedings*, Volume 1406 of *Lecture Notes in Computer Science*. Springer.

Thompson, E. (1968). The projective theory of relative orientation. *Photogrammetria 23(1)*, 67–75.

Tordoff, B. and D. Murray (2002). Guided sampling and consensus for motion estimation. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen (Eds.), *Computer Vision - ECCV 2002 7th European Conference on Computer Vision, Copenhagen, Denmark, May 28-31, 2002. Proceedings, Part I*, Volume 2350 of *Lecture Notes in Computer Science*. Springer.

Torr, P. and A. Zisserman (1997). Robust parameterization and computation of the trifocal tensor. *Image and Vision Computing 15*, 591–605.

Überhuber, C. (1997). *Numerical computation – methods, software, and analysis*. Berlin: Springer.

Vieville, T. and Q. Luong (1993). Motion of points and lines in the uncalibrated case, Report RR-2054, INRIA.

von Sanden, H. (1908). *Die Bestimmung der Kernpunkte in der Photogrammetrie*. Ph. D. thesis, Göttingen University.

Weinshall, D., M. Werman, and A. Shashua (1996). Duality of multi-point and multi-frame geometry: Fundamental shape matrices and tensors. In B. Buxton and R.Cipolla (Eds.), *Computer Vision - ECCV'96, 4th European Conference on Computer Vision, Cambridge, UK, April 15-18, 1996, Proceedings, Volume II*, Volume 1065 of *Lecture Notes in Computer Science*, pp. 217–227. Springer.

Weng, J., T. S. Huang, and N. Ahuja (1992). Motion and structure from line correspondences: closed-form solution, uniqueness, and optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14(3)*, 318–336.