

Design and Development of a Calibration Solution Feasible for Series Production of Cameras for Video-Based Driver-Assistant Systems

A thesis accepted by the Faculty of Aerospace Engineering and Geodesy of the University of Stuttgart
in partial fulfilment of the requirements for the degree of Doctor of Engineering Sciences (Dr.-Ing.)

by

Mehrdad Nekouei Shahraki

Born in Iran

main referee: apl. Prof. Dr. Norbert Haala

co-referee: Prof. Dr. Uwe Sörgel

co-referee: Prof. Dr. Hans-Gerd Maas

date of defence: 20.11.2020

Institute for Photogrammetry

University of Stuttgart

2022

Abstract

To ensure making valid decisions with high accuracy in machine vision systems such as the video-based driver assistant systems, one of the primary key factors is to have accurate data and measurements. This implies that for these systems, we need accurate camera calibration. Since these cameras cover a wide range of optical designs there is a need to be able to model their optical characteristics. It is necessary to estimate accurate camera calibration for various applications.

Conventional and classical camera calibration methods have specific limitations such as having time-consuming real-time calculation for image rectification. This feature is especially needed in mobile embedded systems, which are processing real-time data on low-power processing platforms. These systems require a very fast approach that takes about one hundredth of a second to analyse the calibration data and correct the image observations in real-time on high frame rates e.g. 30 fps. The correction thereby needs to comply with hardware specification installed in these systems including platforms with low thermal design power.

Recently, low-cost objectives are more often required in cost-effective systems and solutions to target growing markets. They also need to comply with the high robustness standards of the automotive industry. Therefore, these lenses have a relatively more complicated distortion, or local distortion as a result of their cost-effective structural design and construction. Classical models have limitations in modelling the local lens distortions. This has - to some extents - been improved by the recently introduced models such as the Legendre polynomials from Tang by using the high-order parameters of the bivariate polynomial. The need to further improve the accuracy of local lens distortion modelling beyond the current models, and in the same time achieve a very high real-time rectification capability, implies the necessity to introduce new solutions.

We introduce the techniques for performing the intrinsic camera calibration and modelling the camera lens distortion with good accuracies, while yet allowing real-time calculation possibilities to perform image distortion and undistortion. The concept is based on free-function modelling in an a posteriori calibration step using the initial distortion estimation and the corresponding residuals on the observations as input information.

Free-function modelling is the technique of numerically and locally fitting functions using discrete expansion to the camera distortion field. The idea behind free-function modelling is that instead of finding the best calibration model or mathematically the best function to fit to our observation data, we assume unknown functions in our calibration model or in other words the calibration model to be analytically unknown and therefore remove explicit constraints in the model. This implies that the functions fit themselves to the observation data no matter what kind of mathematical form they possess.

On one hand, this increases the flexibility of the distortion model to fit itself to different lens distortions and optical designs and furthermore to be able to model the very local lens distortions. On the other hand, based on the pure mathematical nature of this model, a physical interpretation of the evaluated parameters is difficult. However, the flexibility of this model is much beneficial in cases when we mostly focus at modelling accuracies. The correlation between camera extrinsic and intrinsic parameters is not higher than classical methods as a physically-motivated calibration model is estimated initially.

Using the free-function model one can normally observe enhancements in accuracy - in comparison with classical models – especially in the modelling of the local lens distortion. We can see accuracy increase in the residuals of the modelling, whilst still maintaining a faster processing time for recalculation of the distortion model. This improvement depends on the complexity of existing distortion and the complexity and order of the free functions, e.g. the number of function series.

The quality of local distortion modelling depends on the number and distribution of the observed points and the initial accuracy of point coordinates in object-and-image space i.e. point detection accuracy. Therefore by increasing the number of control points and improving their distribution the quality of lens modelling would be improved; a characteristic which is predominantly not present in the classical methods.

Kurzfassung

Um mit Hilfe des maschinellen Sehens, welches beispielsweise in video-basierten Fahrerassistenzsystemen eingesetzt wird, korrekte Entscheidungen treffen zu können, müssen die Messsysteme Daten mit hoher Reliabilität liefern. Eine Schlussfolgerung daraus ist, dass die verwendeten Kameras sehr genau kalibriert werden müssen. Da die optischen Eigenschaften von Kameras sehr stark variieren können, ist es notwendig flexibel anpassbare Modelle zu entwickeln, mit denen das Abbildungsverhalten der Kameras mit hoher Genauigkeit beschrieben werden kann.

Konventionelle und klassische Kalibriermethoden für Kameras haben spezifische Limitierungen, beispielweise hinsichtlich der Echtzeitfähigkeit für die Bildrektifizierung. Diese Eigenschaft ist besonders in Mobile eingebettete Systeme erforderlich, welche die Echtzeitdaten auf Verarbeitungsplattformen mit niedriger Leistung und niedrigem Energieverbrauch bearbeiten. Diese Systeme erfordern Modelle, welche es ermöglichen, mit Berechnungszeiten unterhalb von 10 ms Bilder zu rektifizieren und zu analysieren. Dies ist nötig, um bei Bildraten von 30 fps die Bilder in Echtzeit zu verarbeiten. Die Korrekturberechnung muss daher die Hardware-Spezifikationen einhalten, die oftmals durch niedrige Leistung und thermische Auslegung gekennzeichnet sind.

Die wachsenden Märkte im Bereich Fahrerassistenz Systeme führen dazu, dass vermehrt robuste und kostengünstige optische Module in Kameras verwendet werden, um rentable Systeme anbieten zu können. Diese Module müssen die hohen Robustheitsstandards und Normen der Automobilindustrie erfüllen und gleichzeitig passende Kosten in die Serienproduktion haben. Deshalb haben dort verbaute Linsen eine verhältnismäßig komplizierte lokale Verzeichnung infolge ihres zuverlässigen und kostengünstigen Aufbaus. Klassische Modelle haben Limitierungen in der Modellierung der lokalen Verzeichnungen. Das wurde, in gewissem Maße, durch kürzlich eingeführte Modelle wie die Legendre-Polynome (Tang, et al., 2012 A) durch das Verwenden der Parameter der hohen Ordnung des zweidimensionalen Polynoms verbessert. Das Bedürfnis, die Genauigkeit der lokalen Linsenverzeichnung weiter zu verbessern und gleichzeitig eine sehr hohe Echtzeitkorrektur-Fähigkeit zu erreichen, deutet auf die Notwendigkeit hin, neue Lösungen einzuführen.

In der vorliegenden Arbeit werden neue Techniken eingeführt, um die intrinsische Kamerakalibrierung, Bildverzerrung und Bildrektifizierung mit hohen Genauigkeiten durchzuführen. Diese Techniken erfüllen gleichermaßen die Anforderungen an die Modellierung der lokalen Verzeichnung, als auch an die Möglichkeiten zur Echtzeitberechnung. Das Konzept basiert auf der sogenannten freien Funktion, die in einem A-posteriori-Kalibrierungsschritt, mit der initialen Evaluierung der Verzeichnung und den entsprechenden Residuen auf den Beobachtungen als Eingangsinformation, die Modellierung durchführen kann.

Unter dem Begriff 'freie Funktion' wird hier die Technik verstanden, lokale Funktionen mit der diskreten Erweiterung am Kameraverzeichnungsfeld zu evaluieren. Die Idee hinter dem Modellieren der freien Funktion besteht darin, dass wir analytisch-unbekannte Funktionen in unserem Kalibrierungsmodell annehmen. Anstatt ein physikalisches Modell zu finden, welches das Abbildungsverhalten der Optik möglichst gut beschreibt, wird hier das Kalibrierungsmodell als analytisch bzw. algebraisch unbekannt angenommen (non-Elementar). Das bedeutet, dass die Funktionen sich den Beobachtungsdaten anpassen, unabhängig davon, welche mathematische Form sie besitzen. Einerseits erhöht das die Flexibilität des Verzeichnungsmodells, um sich verschiedenen Linsenverzeichnungen und optischen Designs anzupassen. Darüber hinaus hilft dies, die lokalen Linsenverzeichnungen zu modellieren. Andererseits, gegründet auf der rein mathematischen Natur dieses Modells, ist eine physikalische Interpretation der bewerteten Parameter schwierig. Jedoch ist die Flexibilität dieses Modells in vielen Fällen nützlich, wenn wir uns größtenteils an Modellierungsgenauigkeiten orientieren. Die Korrelation zwischen den extrinsischen und intrinsischen Kalibrierparametern ist nicht höher als in klassischen Kalibriermethoden, weil zu Beginn ein physikalisch motiviertes Kalibriermodell verwendet wird.

Mit dem Modell der Freien Funktion kann man normalerweise Verbesserungen in der Genauigkeit - im Vergleich mit klassischen Modellen - besonders im Modellieren der lokalen Linsenverzeichnung beobachten. Wir können eine Genauigkeitserhöhung in den Residuen des Modellierens sehen, während wir zusätzlich eine schnellere Verarbeitungszeit für die Neuberechnung des Verzeichnungsmodells aufrechterhalten. Die Verbesserung hängt von der Komplexität der vorhandenen Verzeichnung und der Komplexität und Ordnung der freien Funktionen, z.B. die Zahl der Funktionsreihe ab.

Die Qualität des Modellierens der lokalen Verzeichnung hängt von der Zahl und der Verteilung der beobachteten Punkte an, und ist zusätzlich an der anfänglichen Genauigkeit von Punkt-Koordinaten im Welt-und-Bildraum und die Punkt-Entdeckungsgenauigkeit abhängig. Deshalb wird durch das Steigern der Kontrollpunktanzahl und durch eine bessere Verteilung die Qualität der Linsenmodellierung verbessert; eine Eigenschaft, die in den klassischen Methoden überwiegend nicht gewährleistet ist, ohne den Freiheitsgrad des Modells zu ändern.

Contents

Abstract	3
Kurzfassung	5
1. Introduction	10
1.1. Preface	10
1.2. Motivation	11
1.3. Outline	13
2. Theory and Background on Camera Calibration	14
2.1. Camera Optical Structure	14
2.2. Close-Range Photogrammetry	17
2.2.1. Stereographic 3D Modelling and Reconstruction	18
2.2.2. Structure-from-Motion and Multi-View Stereo	19
2.3. Video-Based Driver Assistant Systems	20
2.4. Camera Calibration and Calibration Models	25
2.4.1. Geometry of Perspective Projection	26
2.4.2. Geometry of Omni-Directional Projection	27
2.4.3. Calibration Distortion Models	31
2.5. Conventional Calibration Solutions and Methods	40
2.5.1. Collimator Calibration (Light-Ray Measurement)	40
2.5.2. Bundle-Block Calibration	41
2.5.3. Test Stand Single-Shot Calibration	44
2.5.4. Other Calibration Solutions	46
2.6. Image Marker Measurements and Estimation	47
2.6.1. Contour-Track Marker Detection	48
2.6.2. Sub-Pixel Pattern Cross-Correlation	49
2.6.3. Ellipse Fitting Marker Detection	51
3. Extended Omnidirectional Camera Model	52
3.1. Main Omnidirectional Projection Model	53
3.1.1. Internal and External Lens Modelling (Projections)	53
3.1.2. Projection Coupling	54
3.2. Direct Transformation: World to Camera (W2C)	55
3.3. Linearization (Taylor series derivatives)	56
3.4. Setting-Up Equations System	60
3.4.1. Geometrical parameter correlation	60
3.4.2. Distortion model for extended fisheye calibration	61
3.5. Inverse Transformation: Camera to World (C2W)	62
3.6. Extracting Physical Parameters and Limits	65

3.7.	Generalized Projection and Runtime Optimization.....	66
3.7.1.	General Angular W2C Projection	66
3.7.2.	General Angular C2W Projection	67
3.7.3.	Runtime optimization of C2W transformation	69
4.	The Free-Function Calibration Distortion Model	70
4.1.	The Mathematics of the Free-Function-Model.....	72
4.1.1.	Function Superposition	73
4.1.2.	Free-Function Model for Distortion Estimation.....	74
4.1.3.	Free-Function Model for Undistortion Estimation.....	75
4.1.4.	Discrete Expansion Equivalency	76
4.2.	Implementation, Real-Time Calculation and Optimization	77
4.2.1.	Performance Evaluation	80
4.3.	Modelling the Local Lens Distortions.....	82
4.4.	Calibration Strategy and Process Flow	84
4.4.1.	Distortion Model and Initial Calibration	84
4.4.2.	Updating the Distortion Field and Local Distortion.....	84
4.4.3.	Using Free-Function Calibration Model.....	84
4.4.4.	Finalizing the Calibration Process	85
5.	General Requirements for Calibration Test Stands	86
5.1.1.1.	Maintaining Calibration Accuracy for Complete Image Area.....	86
5.1.1.2.	Additional Demands on Optical Tests	86
5.1.1.3.	High Repetitive Accuracy and Automatic Calibration	87
6.	Hardware Requirements and Test Stand Design	88
6.1.	Specifications of Cameras under Calibration.....	88
6.2.	Pre-Analysis and Test Stand Design Criteria.....	90
6.2.1.	Pre-Analysis: Test Stand Point Distribution	90
6.2.2.	Pre-Analysis: 2D Image Measurement Accuracy	92
6.2.3.	Pre-Analysis: Test Stand 3D Depth Configuration.....	93
6.2.4.	Pre-Analysis: 3D Test Stand Measurement Accuracy	97
6.3.	Test Stand Specifications for Perspective Camera	98
6.3.1.	Test Stand Design Details.....	98
6.4.	Test Stand Specifications for Omni-Directional Camera.....	100
6.4.1.	Test Stand Design Details.....	100
6.4.2.	Design Statistical Analysis and Point Distribution.....	103
6.5.	Test Stand calibration of Perspective Camera	104
6.6.	Test Stand calibration of Omnidirectional Camera.....	107
6.7.	Discussion on the Values of Repetitive Accuracy.....	109
7.	Performing Test Stand Calibration using Free-Function Model	110
7.1.	Free-Function Calibration on Perspective Camera	110
7.1.1.	Accuracy and Error Analysis	112
7.2.	Performing Free-Function Calibration on Omnidirectional Camera	115
7.2.1.	Accuracy and Error Analysis	117
7.3.	Further Analysis of Lens Distortion Modelling using FF-Model on Real-World Data ..	120
7.4.	Discussion on Accuracy of Free-Function Model.....	124
8.	Summary and Conclusion	125
9.	Outlook.....	127

10. Appendix: Statistical Model and Error Analysis	128
10.1. Parameter Estimation System and Algorithms	128
10.2. Confidence Interval Test	130
10.3. Correlation, Independency and Significance	131
11. Appendix: Pre-Analysis and Network Design	132
12. Appendix: Post-Analysis and Network Design Evaluation.....	133
13. Appendix: Evaluation of Mathematical Models for Analysing Fisheye Distortion ..	134
14. Appendix: Abbreviations	139
15. Appendix: Conventions.....	140
Bibliography.....	141
Curriculum Vitae	149

*“Accuracy of observations
is the equivalent of
accuracy of thinking”*

*Wallace Stevens*¹

1. Introduction

1.1. Preface

Accurate observations and measurements are necessary for making accurate decisions. This would give us quality control over our system, which is an important factor for sustaining any system. Since the first cameras were built, and the possibilities for using cameras and images in different fields were discovered, there have been many applications introduced which benefit from use of images. In many cases, a camera, or more precisely an image sensor, is hold as the main source of our observation. Therefore, solutions have been created in order to make sure about the accuracy of these observations, and also to introduce ways in which we can extract useful information from images, that help us further in e.g. making decisions based on these information. Naturally, the accuracy of our decisions is based on the accuracy of our information, and the later depends on many parameters such as the accuracy of our observations and our control over our observation device. The accuracy of our image observations and the world-to-image connection in computer vision applications is satisfied by the means of calibration.

Camera calibration has been studied for quite a long time since the era of photogrammetry started and has been developing alongside the development in camera technology and our processing capabilities. By everyday increasing needs for using images in different fields, there is an increased demand for calibration solutions that fits the always-increasing requirements for accuracy and speed. As the new applications are introduced every day, camera calibration is becoming a major involving field. Continuous improvements and developments in this field are advantageous to any imaging system.

¹ Wallace Stevens (1879-1955) modernist, lawyer and poet

“If observed facts of undoubted accuracy will not fit any of the alternatives it leaves open, the system itself is in need of reconstruction”

*Talcott Parsons*²

1.2. Motivation

With the leap forward in low-power computational processing and the introduction of digital image processing, variety of new applications are being introduced in the field of mobile photogrammetry and machine vision. The recent extensive developments in the automotive industry and the introduction of driver assistant systems is a valid example of this fact. This makes it necessary to ensure about the quality and trustability of our cameras and images integrated in our driver assistant system and thus brings us the need for accurate testing and calibration solutions.

The specifications of the latest driver assistant systems, similar to modern photogrammetric applications such as three dimensional modelling, require very accurate image and world geometric processing and this presumes a good camera calibration for various video-based applications and functions. These are applications such as lane-detection and lane-keeping, autonomous emergency braking, autonomous parking, three-dimensional top-view and surround-view. These functions are benefiting from recent advancements in machine vision, image processing and photogrammetry. In addition, a very fast approach is required to analyse the calibration data and correct the image observations in real-time, which should of course comply with the hardware installed in these systems. The key points for these hardware are extreme reliability, stability and low power consumption using designs based on hardware such as ARM processors, low-power CPUs or customized ISP (Image Signal Processing) modules. This leads to a very different design and performance compared to the current market PC. Besides, most of the above applications running on these platforms are performed in runtime. Therefore, by looking at both sides we understand the need to have a very high accuracy, flexibility and speed for our calibration, to deliver the best results with the applications running on special hardware.

In photogrammetry in most of the real-time applications, modelling of the camera optics is generally divided in two parts. The first part is the estimation of the optical distortion by considering a mathematical model, which represents the mapping process from world to pixel coordinates. The second part is the rectification of this effect in the image to eliminate the distortion and provide a distortion-free image or apply corrections to the measured image coordinates. This simplifies the further image processing steps using ideal perspective transformation. The above discretion should always be considered to be able to discriminate between the two different operations and their corresponding applications regarding camera calibration. In digital photogrammetry we determine the distortion by estimating an analytical distortion model with which it is possible to generate a distortion-free image.

² Talcott Parsons (1902-1979) Sociologist

In classical or analogue photogrammetry, for certain applications like bundle-block adjustment or structure-from-motion, we measure the image coordinates, consider the distortion correction in the equations for optical distortion and compute the 3D coordinates while simultaneously modelling the optical distortion. However, these calculations are usually performed offline, on powerful systems. These heavy processes are not suitable for real-time image rectification of a stream of frames in low-power real-time systems.

The classical camera calibration and image rectification solutions in photogrammetry - as mentioned above - are usually designed for and performed on platforms with high processing power. In the driver-assistant systems there are systems which offer a processing power less than an average market PC processor. Recently there has been much advancement in the field of camera calibration by introducing different models, increasing accuracy and internal parameter stability. But still most of the today's calibration solutions and calibration models which deliver acceptable accuracies are not suitable for runtime applications. They usually have a lot of calculation per-pixel that takes time to perform on low-power processors. This fact should be considered when designing and developing solutions for these special hardware. Thus, we need to design a calibration solution that fits to our hardware in order to achieve this goal.

The calibration model that we need, should be capable of having desired sub-pixel accuracy comparable to today's calibration models or better, and should be very fast for realizing a real-time and optimizable rectification approach with a high calculation speed. This is critical in processing and rectifying multi-frames per second on a system with special hardware specifications. As an example for speed requirement, it could be mentioned that these systems process image sequences with at least thirty frames per second, and most of the performance of these systems is dedicated to the complex algorithms running online to process the images and handle the dataflow. Therefore, although the intrinsic calibration should be very accurate on these systems, it also has to be fast in recalculation for a high frame rate.

1.3. Outline

The first chapter provided a brief introduction for this work. The second chapter, we read through some theoretical backgrounds on close-range photogrammetry and different applications in this field such as video-based driver-assistant systems. Then we study camera calibration in more details, mainstream and also newer calibration models and their mathematical calibration projection and distortion models. After that we follow by reviewing some of the conventional calibration solutions and also several image measurements techniques used in those calibration solutions.

In the third chapter we develop an extended omnidirectional camera model based on the work of Christopher Mei and further expand it in details to get transformation functions in a classical photogrammetric form. Then we do a detailed calculation on the generalized projection function and follow this to achieve the solutions for runtime optimizations.

The fourth chapter introduces the free-function calibration model. This model has been developed to have the capabilities of high accuracy distortion modelling to be able to model the local lens distortion, and at the same time have a high runtime recalculation performance.

In chapter five, we review the general requirements that we consider when designing new calibration test stands, such as high repetitive accuracy and additional optical testing capabilities. In chapter six, we take these requirements into account and design two new test stands for a perspective and a fisheye camera using pre-analysis calculation to determine design criterions based on the former requirements. After that, we perform camera calibration using these new test stands to get a sense of the quality of parameter reproducibility which has been set as one of the main design goals.

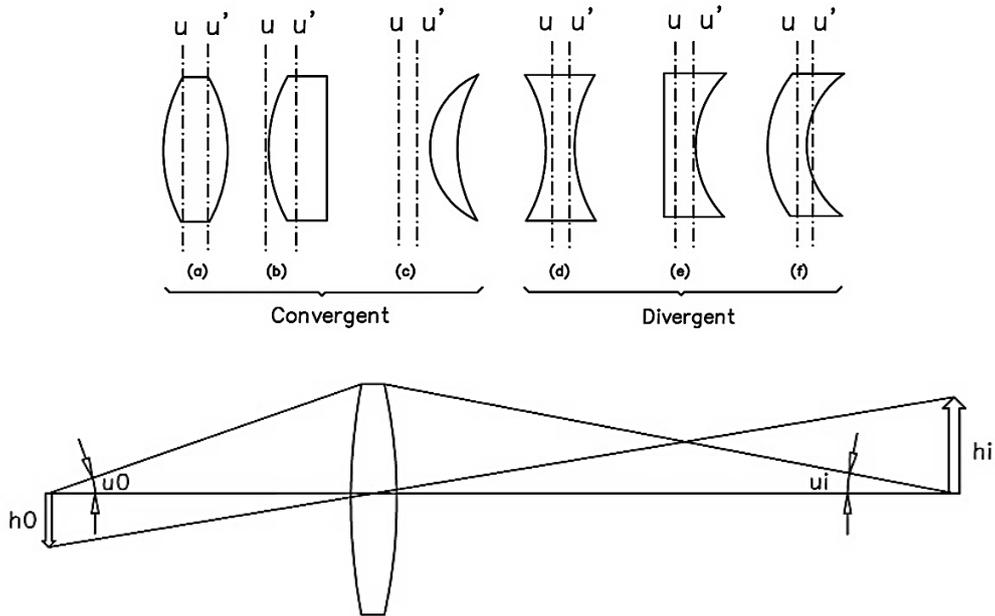
Finally in chapter seven, we continue the work on the previous chapters by performing the free-function calibration using the newly designed test stands and analyse the calibration quality and the accuracy results. The summary and conclusion of this work is provided in chapter eight. Furthermore, complimentary information is available in appendixes from chapter ten to fifteen.

2. Theory and Background on Camera Calibration

The general principles and the basics of our work are studied in this chapter. First, we review the general optical structure of the camera. Then we discuss the theoretical fundamentals and applications in close-range photogrammetry. After that, we introduce the principals of video-based driver-assistant systems, different cameras used in them and their variety of applications. Next, we study the different camera calibration methods. Finally, we go through different image measurement techniques used in close-range photogrammetry and camera calibration solutions.

2.1. Camera Optical Structure

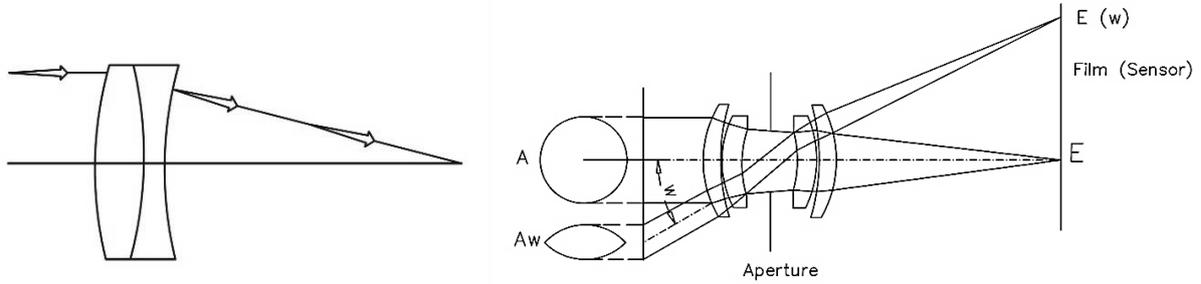
We need to study the geometrical features of the camera optical module first, to be able to understand their details and use them effectively. This helps us to understand accurate analytical relations between the image space and object space. These includes basic features such as the equivalent focal length and the optical distortion effects which are mostly based on the camera optical structure. There has been recent works on all-reflective cameras, which use the concept of reflection (e.g. mirrors) instead of refraction (e.g. lenses) as they are beneficial in many fields of science (Richter, et al., 2013). In this work, we focus on refractive cameras with the optical modules based on lens-combinations. The camera optic is usually made up of a combination of different lenses with which an ideal projection could be achieved.



a Fig. 2.1. Example of design of principal optical elements
b a: Common types of lenses: (a) double-convex; (b) plano-convex; (c) convergent meniscus; (d) double-concave; (e) plano-concave; (f) divergent meniscus. u and u' are the unit planes, light being assumed to be incident from the left. Illustrated after (Born & Wolf, 1999) Fig. 4.15.
 b: Imaging by a simple lens, The object height is h_0 and the image height is h_i , with u_0 and u_i being the corresponding slope angles, illustrated after (Bass, et al., 1995) p. 1.5.

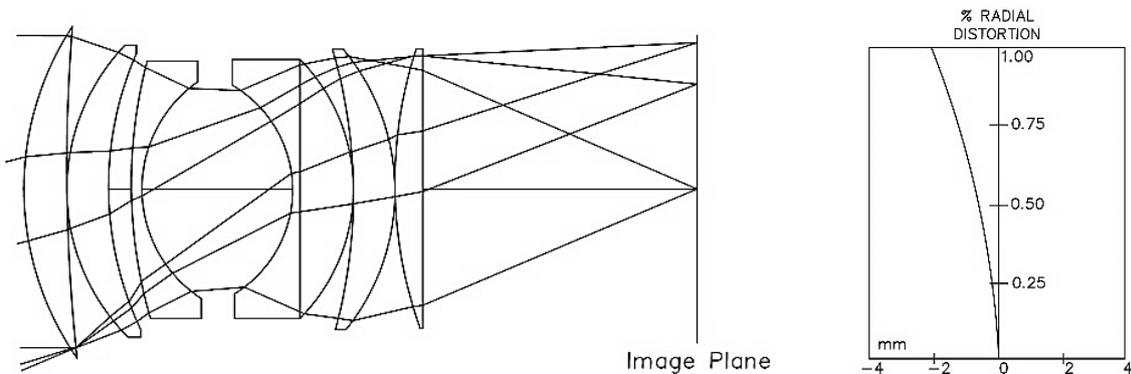
Fig. 2.1 illustrates some of the fundamental optical elements used in current cameras, and their relevant geometrical design parameters. We can see single lens designs with their defined parameter.

The knowledge about the most important parameters of a lens such as the curvature radius, diameter and the effective focal length enable us to combine them to create and realize special optical behaviour. Primary lens elements such as plano-convex and plano-concave, are used together to form complex combined lens systems. This enables us to achieve certain optical properties such as different optical projection specifications.



a **b** Fig. 2.2. Illustrating an example of simple lens combinations
 a: Typical achromatic doublet lens, illustrated after (Bass, et al., 1995) p. 1.20, 1.23
 b: Formation and reduction of entrance pupil in a lens-combination based on the angle w . Illustrated after (Schröder & Treiber, 2002).

A schematic example for the combination of the lens elements is illustrated in Fig. 2.2. Using these single lens elements, one can design optical complex modules with different characteristics to create standard wide-angle perspective cameras or fisheye cameras. Most of the today's cameras used in photogrammetric and computer-vision applications are constructed using similar lens elements.



a **b** Fig. 2.3. Illustrating an example of a complex lens combination
 a: An example of a perspective camera (55-mm F/1.2 for 35-mm SLR)
 b: The radial distortion of the same lens construction (left)
 Illustrated after (Bass, et al., 1995) p. 16.3.

Different features and parameters of the optical module such as depth of focus, circle of confusion, f-stop, focusing distance and depth of field are the product or requirement of a special lens design. As an example, the radial distortion displayed in Fig. 2.3.b is an optical feature which could be described as a deviation of the optical projection from the ideal perspective projection. This topic is further discussed in chapter 2.4.

Using the lens combination techniques, one can create cameras with variety of optical projection behaviours such as very big opening angles (field-of-view). These cameras, which are known as fisheye cameras are used in many computer-vision and photogrammetric applications. Benefiting from the very big opening angle, they provide large amount of information from the near environment in the image.

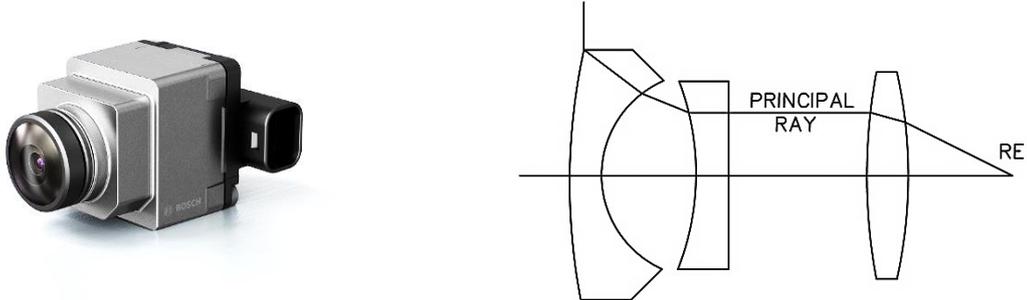


Fig. 2.4. Examples of fisheye camera objectives

a: A fisheye camera used in surround-view driver assistant systems (Courtesy of Robert Bosch GmbH)

b: Illustration of a simple fisheye camera structure (inverse galilean with hemispheric field of view)

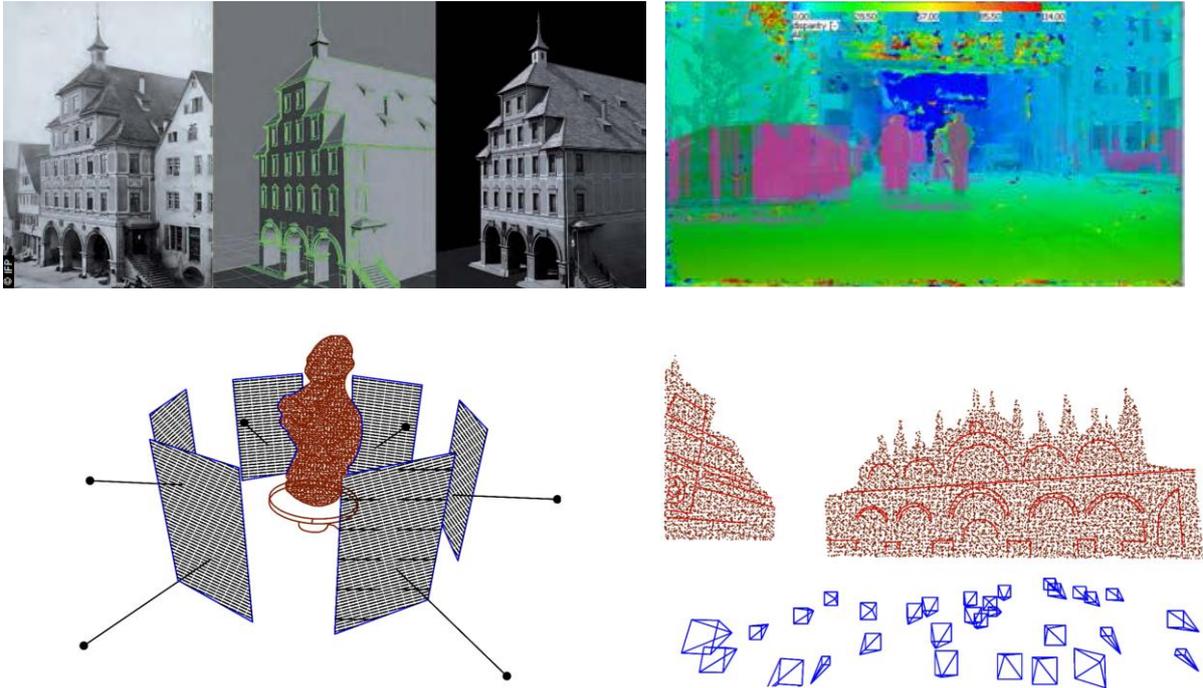
Illustrated after (Bass, et al., 1995) p. 2.15.

As it is illustrated in Fig. 2.4 one can see that the design of fisheye lenses is based on combination of the very similar principal lens elements used in central-projection cameras. This makes it possible to realize different overall optical projection behaviour in our camera such as extreme opening angles of bigger than 190 [deg]. This is very beneficial especially in close-range photogrammetric applications where one can cover the whole 360 [deg] of the surrounding environment using at least two cameras.

In this chapter we have done a short review on the principal structure of different cameras. Using cameras with different optical projections enables us to take images with different geometry which are suitable in different applications in the field of photogrammetry and computer-vision. In the next chapter we discuss the basics of close-range photogrammetry and some of the applications in this field.

2.2. Close-Range Photogrammetry

“Photogrammetry encompasses methods of image measurement and interpretation in order to derive the shape and location of an object from one or more photographs of that object. In principle, photogrammetric methods can be applied in any situation where the object to be measured can be photographically recorded.” (Luhmann, et al., 2011). There are many applications in the field of photogrammetry and close-range photogrammetry these days such as architectural heritage modelling, indoor modelling, video-based navigation systems, and optical diagnostics and so on. These applications benefit from the recent works in this field. In all of them we deliver the geometric information of the object space by processing the images taken from them.



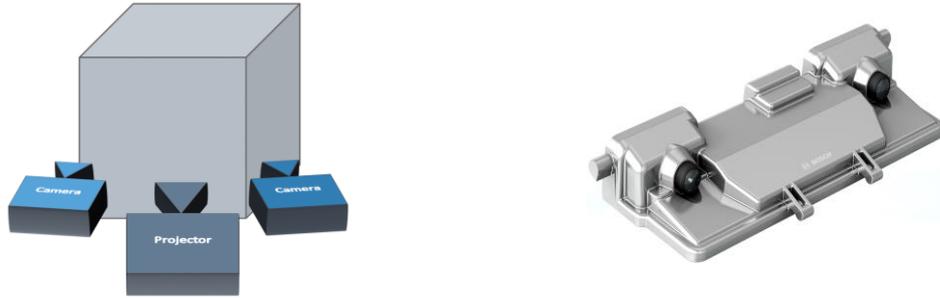
a	b	Fig. 2.5. Example of different applications in Close-Range photogrammetry
c	d	

a: 3D modelling of Architectural heritages using close-range photogrammetry (Courtesy of ifp, University of Stuttgart)
 b: Real-time street modelling for vehicle trajectory control via real-time Stereo 3D modelling. Vehicles, street surface, pedestrians and buildings are color-coded (Courtesy of Robert Bosch GmbH) c: Image-Based 3D reconstruction, Illustrated after (Hernández & Furukawa, 2013). d: SfM model of San Marco Square in Venice, Illustrated after (Hernández & Furukawa, 2013). Few of the camera positions are also shown in the reconstructed scene.

In Fig. 2.5, we can see some of the close-range photogrammetry applications used nowadays. The three-dimensional reconstruction of objects such as buildings and architectural heritage sites (Hernández & Furukawa, 2013) using techniques such as dense image matching is a critical topic in this field. This is very important, as generating detailed three-dimensional models from those structures is very beneficial in the context of preserving the cultural heritage and historical sites. The reconstruction of three-dimensional models of objects could be performed in real-time and provide us the necessary information to analyse the surrounding object space, followed by the analysis of the geometric shape, distance and the relative movement of the object with respect to the camera (Fig. 2.5.b).

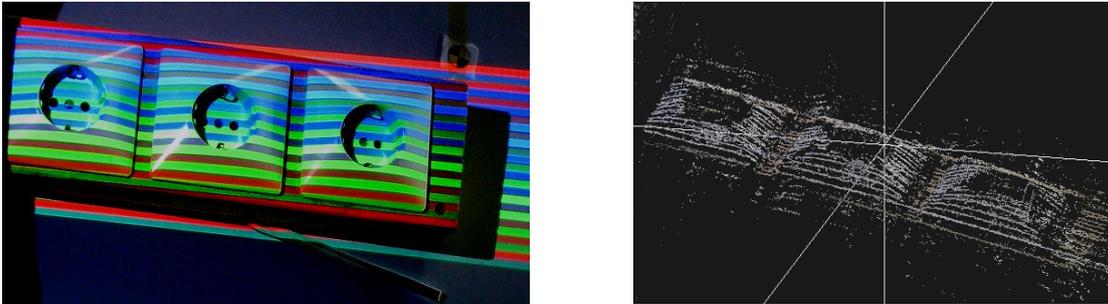
2.2.1. Stereographic 3D Modelling and Reconstruction

Principally, by using two or more cameras and an optional video projection system we can simultaneously capture images from an object, perform stereographic image matching by finding corresponding points in both images and then reconstruct its three-dimensional digital form. The most common setup is illustrated in Fig. 2.6.



a **b** Fig. 2.6. Stereographic 3D modelling in Close-Range photogrammetry
 a: Illustration of a stereo camera setup with two single cameras and a video projector
 b: Picture of a stereo camera, consisting of two single imager modules inside the housing.
 (Courtesy of Robert Bosch GmbH) Further discussions in [2.3].

Three-dimensional reconstruction is a key applications when we want to perform video-based diagnosis of an object to ensure certain geometrical features. The stereo object reconstruction can result into a dense 3D point cloud, and makes it possible to create a virtual three-dimensional model based on the generated point-clouds.

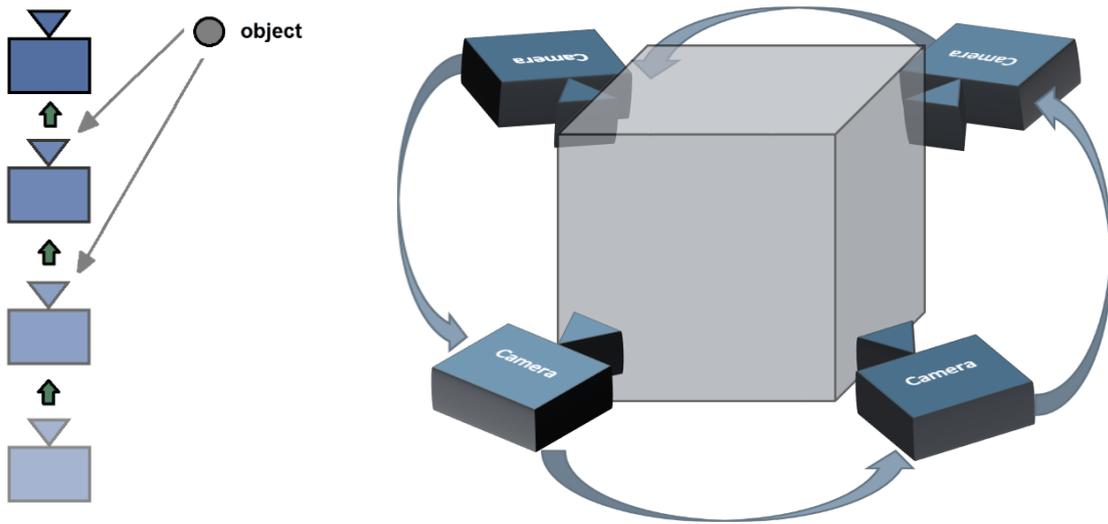


a **b** Fig. 2.7. Example of different applications in Close-Range photogrammetry
 a: Pattern projection and stereo imaging based on (Mazaheri & Momeni, 2008) using stereo matching LCS algorithm (Apostolico, et al., 1992) (Cormen, et al., 2009) for modelling a power strip device b: Point cloud from 3D reconstruction of the power strip device in a, generated from a single image.

The stereo image matching is used on both images to estimate the pixel correspondence and create the 3D point coordinates. Stereo matching is typically based on epipolar geometrical constraint, which implies an accuracy requirement of around 0.1 [pixel]. As a common application, closely related to computer-vision field, one can verify the form of a product which is being measured by inspecting the reconstructed geometrical model of that object without having direct contact as illustrated in Fig. 2.7. This is also an important and necessary task when we consider product quality-control.

2.2.2. Structure-from-Motion and Multi-View Stereo

In photogrammetry, structure-from-motion (SfM) “is the process of reconstructing 3D structure from its projections into a series of images taken from different viewpoints”. (Schönberger & Frahm, 2016) Generally, by having a single camera, one can use the camera’s motion while capturing and calculate the orientation and position in object space for each captured image. After that by performing techniques such as stereo dense matching on consequential images the 3D object reconstruction could follow. This technique is often utilized when we are modelling an object using a single camera and have the possibility to move the camera in certain directions and positions.



a **b** Fig. 2.8. Example of different SfM approaches in Close-Range photogrammetry
 a: Forward motion of the camera for SfM and object reconstruction, vectors of relative positions of object to camera are illustrated.
 b: Camera motion around an object for a complete 3D modelling and geometrical reconstruction Illustrated after (Westboy, et al., 2012).

As illustrated in Fig. 2.8, we can perform the 3D modelling of an object by continuously rotating the camera around it or moving the camera relative to the object. Using the consecutive captured images to perform stereo dense matching and 3D reconstruction, the object’s geometry can be virtually modelled. This is often used in variety of applications where the movement of camera can create a relative translation with respect to an object. For example, a front-facing camera could be mounted in a vehicle, where the vehicle’s forward motion (Fig. 2.8.a) creates relative translation to the surroundings, thus enables us to calculate object positions with respect to the vehicle. These applications are often categorized as video-based driver assistant systems.

The video-based driver-assistant systems are considered as some of the recent and highly developing applications in the field of photogrammetry. Functions such as EBA (Emergency Brake Assist) or AEB (Automatic Emergency Braking) using only video sensors have been recently introduced, and a diversity of new functions are being developed on multiple platforms such as commercial vehicles. They take advantage of variety of algorithms based on close-range photogrammetry techniques. In the next chapter, we discuss the principles of driver-assistant systems and review some examples in this field.

2.3. Video-Based Driver Assistant Systems

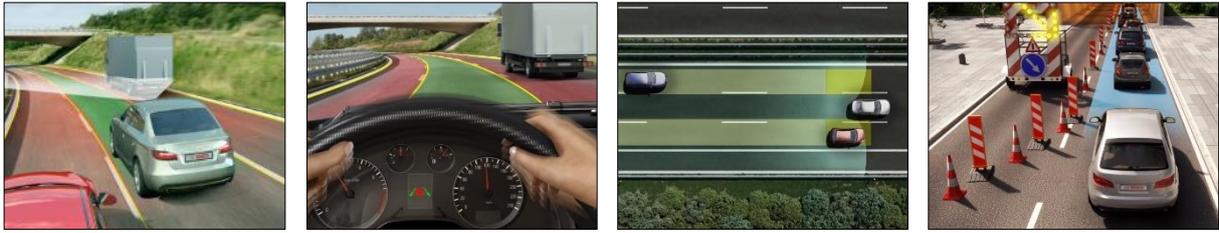
Driver Assistant (DA) systems are systems which gather data related to real-time driving safety from the vehicle and its surroundings, provide the driver with the necessary information and aid the driver in critical situations to avoid accidents or if unavoidable, to decrease injuries and damage.

In order to fulfil these goals, the driver assistant systems make use of recent technologies such as intelligent video cameras, radar sensors, LiDAR, ultrasonic and state of the art online data-fusion and data processing techniques. This makes it possible to monitor the environment around the vehicle, derive critical information and make decisions based on a state of situational awareness.



a	b	Fig. 2.9 Variety of Driver-Assistant Cameras a: Front-View mono camera (MPC) b: Front-View stereo camera (SVC) c: Near-Range Surround-View Camera (NRC) d: Night-View Camera (NV) (Courtesy of Robert Bosch GmbH)
c	d	

Since the introduction of driver assistant systems, widespread possibilities have been explored for taking advantage of cameras such as mono/stereo front-view cameras, omnidirectional surround-view cameras as illustrated in Fig. 2.9 . These systems are used for vehicle drive-time trajectory control, situational awareness, crash avoidance, three-dimensional space representation, low-speed manoeuvre assistance. Also the near-infrared (night-view) cameras are often used for night time applications by e.g. increasing and improving the vision quality and sight range in dim-light conditions.



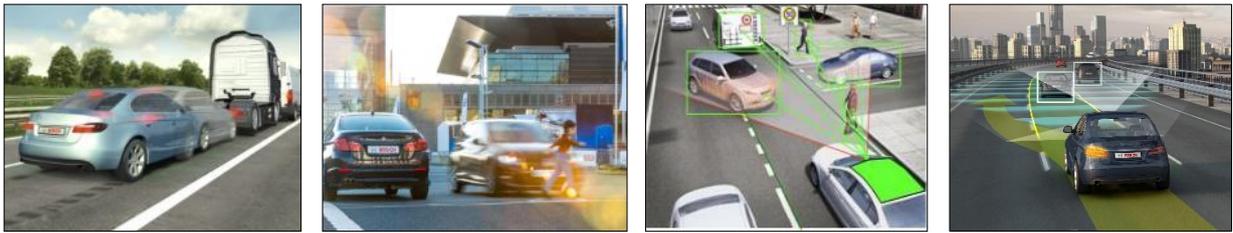
a	b	c	d
---	---	---	---

Fig. 2.10. Examples of Driver-Assistant Applications

a, b: Lane Keeping Assist c: Lane Change, Lane Departure Control

d: Lane Keeping Narrow Zone / Construction (Courtesy of Robert Bosch GmbH)

Applications such as lane keeping assist and the similar related functions are based on detecting the lanes, the surrounding vehicles and using all the data in an online vehicle drive system as illustrated in Fig. 2.10. Employing front-view mono or stereo cameras and using real-time lane detection algorithms together with real-time steering control usually realize these functions. The data are then used to monitor the current situation, steer and keep the vehicle well inside the lanes. These information could be used together with other sensors such as side radar to warn the driver for possible upcoming accidents when attempting to change the lanes or control the vehicle if necessary.



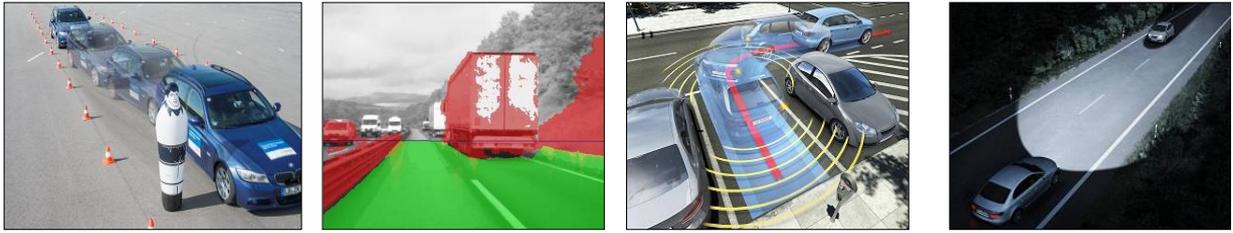
a	b	c	d
---	---	---	---

Fig. 2.11. Examples of Driver-Assistant Applications

a, b: Automatic Emergency Braking

c: Pedestrian Protection d: Highway Assistant (Courtesy of Robert Bosch GmbH)

Another application in driver-assistant systems, as we can see in Fig. 2.11, is the automatic emergency braking which is generally based on monitoring all the vehicles and pedestrians in front of the vehicle and predicting the upcoming situation and deciding on either performing full break or taking other measures to ensure the safety of pedestrians and passengers. These functions are taking advantage of front-view stereo-cameras or radar sensors. For example, the video cameras, depending on the defined scenarios and vehicle speed limits for the corresponding functions, would require certain object-space accuracies. For distances of around 20 [m] a distance accuracy of around 0.5 [m] is required, and an accuracy of 30 [mm] for distances around 5 [m]. As also discussed in chapter 2.2.1 this requirement corresponds to around 0.1 [pixel] in image space for certain types of cameras. This accuracy is required to perform a stable tracking of an object in 3D space, which includes measuring its heading and position relative to the moving vehicle with enough certainty to trigger functions such as emergency braking or use adaptive cruise control with fixed distance to the vehicle in front. The accuracy requirements and how they are derived is discussed further in chapter 6.2.1.



a	b	c	d
---	---	---	---

 Fig. 2.12. Examples of Driver-Assistant Applications
 a: Manoeuvre Assist (Accident avoidance) b: Manoeuvre Assist (Stereo/Disparity-Based)
 c: Automated Parking (Surround View) d: Dynamic Headlight Control
 (Courtesy of Robert Bosch GmbH)

There are many applications and functions in driver assistant systems which are using technologies based on video or other sensors. As illustrated in Fig. 2.12, we can further mention some main functions such as manoeuvre assistant systems based on stereo 3D object reconstruction and road free-space recognition, adaptive cruise control systems using stereo cameras or radar sensors, and dynamic headlight control which are realized by real-time analysis of data from video sensors or from multiple different sources using sensor data fusion techniques.



a	b	c
---	---	---

 Fig. 2.13. Examples of Driver-Assistant Applications
 a: 3DTV: top and trailer view b: 3DSV: top and rear-to-front view
 c: 3DTV: top and rear view (Courtesy of Robert Bosch GmbH)

By using surround-view cameras as illustrated in Fig. 2.13 we can reconstruct and present the surroundings of the vehicle in a live three dimensional surround-view system (3DSV) or using them as online trajectory guide for manoeuvre/park assistant systems. These cameras are fisheye cameras with special optical geometry and extreme opening angle which together can cover the 360 degree surroundings of vehicle. Assuming a relative movement between the vehicle and the surrounding objects, it is possible to use techniques for three dimensional reconstructions of the surroundings, from which we can name the real-time structure from motion (SFM) and dense multi-view stereo techniques [2.2.2]. Alternatively, ultrasonic or radar sensors can provide the system with initial 3D object information that can be used to identify or detect those objects visually in the image if the relative movement between those objects and the vehicle is very low or not adequate. These techniques would help us to achieve situational awareness in many different driving and traffic scenarios and provide us vital geometric and depth information of the vehicle's surroundings.



a	b	c
---	---	---

 Fig. 2.14. Examples of Driver-Assistant Applications

a: 3DTV and park assistant systems b: 3DTV and park assistant systems

c: 3DSV and park and manoeuvre assistant systems

(Courtesy of Robert Bosch GmbH)

As it can be seen in Fig. 2.14, the surround-view cameras are used to create a three dimensional top-view model (3DTV). This system aids the driver in many city traffic situations and are essential in realizing many manoeuvre/park assistant functions. For example we can name functions such as automated parking, assisted parking with trailer-hitch view or panorama view and reverse steering assistant when connected to an extension vehicle.



a	B	c
---	---	---

 Fig. 2.15. Examples of sensor Data-Fusion for Driver-Assistant Applications

a: Data-Fusion: front-view video, surround-view video, radar and ultrasonic

b: Data-Fusion: video and radar with online monitoring

c: Data-Fusion together with provided localized information

(Courtesy of Robert Bosch GmbH)

There are many possibilities for incorporating data from different sensors in a unified system, i.e. sensor data fusion, as shown in Fig. 2.15. For example, using data from ultrasonic sensors, the information could be combined together with the surround-view fisheye cameras to realize functions such as automatic parking. The data could also be treated as additional information in manoeuvre assistant systems. These functions are one of the recent and newest applications offered by driver assistant systems.

As discussed in this chapter, the driver assistant systems are designed to take the necessary measures in critical situations in order to avoid any upcoming accident or decrease the chance of any injury or damage. In these systems, in order to be able to make valid decisions, we require data from variety of optical sensors. The cameras used in these systems cover a big range of optical designs which suit to different major applications and platforms such as different types of ultra-wide-angle fisheye and wide-angle central projection cameras. There are varieties of fisheye cameras being used in driver assistant systems for applications that benefit from a wide scene and higher content information. This is achieved e.g. by specific complex lens combinations which make it possible to have opening angles of more than 180 degrees. This principally gives more information about the surrounding areas of a vehicle and is advantageous for realizing a state of situational awareness in different driver-assistant applications.

The algorithms and techniques used in driver assistant systems (such as Stereo Reconstruction, SFM, 3DSV, 3DTV and etc.) process the images and perform image measurements with high accuracy and sampling depth. Therefore these algorithms require an accurate world-to-camera transformation, camera modelling and therefore an accurate camera calibration. This implies that for the video-based driver assistant systems we need accurate camera calibration as one of the primary key factors to ensure the high accuracy of object reconstruction which finally leads to the high accuracy and confidence of our final decisions and sustainability of our system.

Furthermore, in these systems, the images should first be corrected for optical effects such as lens distortions. This helps us to use a simple transformation model (such as the perspective projection) in our photogrammetric algorithms and achieve higher processing speed and accuracy for most of the real-time functions. To fulfil these requirement, it is necessary to first design calibration systems (hardware and software) which are capable of accurately calibrate these cameras with sub-pixel accuracy requirements, for example 0.1 [Pixel]. These requirements are derived from special use-cases and are discussed in chapter [5], [6.1]. Second, it is required to design calculation and rectification solutions compatible for our real-time processing system.

The above discussion shows the need for performing calibration by introducing a model which delivers high sub-pixel accuracies and in the meantime could be used in real-time for online image rectification. To understand the available techniques and methods in our hand, in the next chapters we do a short review on camera modelling and calibration solutions. After that we design a new software and hardware solution for camera calibration and real-time rectification.

2.4. Camera Calibration and Calibration Models

In close-range photogrammetry we mostly use an ideal perspective projection to describe the world-to-camera transformation. We use this transformation to perform tasks such as three dimensional modelling, pattern recognition, object tracking and etc. For this purpose the accurate behaviour of lens-sensor should be estimated which means the camera needs to be calibrated.

There have been many calibration methods and models being introduced which vary from mechanical to analytical solutions. These days camera calibration is done mostly analytically and for this purpose many calibration models are used (Brown, 1971) (Moffitt & Mikhail, 1980) (Luhmann, et al., 2011) (Tang, et al., 2012 A) (Tang, et al., 2012 B). Generally the relation between object points (X_p, Y_p, Z_p) and image points (x_b, y_b) is defined as follows:

$$\begin{aligned} x_b &= f(X_p, Y_p, Z_p, X_o, Y_o, Z_o, \omega, \varphi, \kappa) + \Delta x \\ y_b &= g(X_p, Y_p, Z_p, X_o, Y_o, Z_o, \omega, \varphi, \kappa) + \Delta y \end{aligned} \tag{Eq. 2.1}$$

In Eq. 2.1, the world to camera projection functions are indicated as f and g . The parameters (X_o, Y_o, Z_o) are coordinates of camera in the object space coordinate system, and the orientation of the camera is denoted as three angles $(\omega, \varphi, \kappa)$. The terms Δx and Δy are called APs (additional parameters). These parameters would model and compensate the remaining lens distortions usually in image plane.

Analytical camera calibration is performed with additional-parameter (AP) models, usually during the bundle-block calibration in laboratory or in-situ calibration. The calibration process is divided into many steps such as positioning the camera in multiple positions or moving the test-field and capturing images at each angle. Thus we can estimate the additional parameters for the intrinsic calibration by processing the image database in a single-shot or bundle-block equation system together with the APs.

As described above, we need to study the camera projection and the corresponding mathematical models first. Then we can proceed to model the camera projection and describe it mathematically. In this chapter, first we discuss the geometry of perspective and omni-directional projection by reviewing models which describe different projection behaviours. After that, we study different standard distortion models which are commonly used as APs in analytical camera calibration and discuss their advantages and disadvantages.

2.4.1. Geometry of Perspective Projection

The geometrical principle of photogrammetry is the analytical relation between image observations and object information. This is described by the collinearity equation³ (Amini, 2006):

$$x_b - x_h + dx = -f \frac{(X_p - X_o) r_{11} + (Y_p - Y_o) r_{12} + (Z_p - Z_o) r_{13}}{(X_p - X_o) r_{31} + (Y_p - Y_o) r_{32} + (Z_p - Z_o) r_{33}}$$

$$y_b - y_h + dy = -f \frac{(X_p - X_o) r_{21} + (Y_p - Y_o) r_{22} + (Z_p - Z_o) r_{23}}{(X_p - X_o) r_{31} + (Y_p - Y_o) r_{32} + (Z_p - Z_o) r_{33}}$$

$$\mathbf{R}_{\kappa\varphi\omega} \text{ (Rotation Vector)} = \mathbf{R}_\kappa \mathbf{R}_\varphi \mathbf{R}_\omega = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad \text{Eq. 2.2}$$

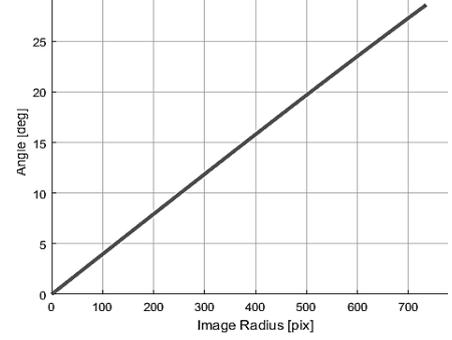


Fig. 2.16. Projection curve of a perspective camera (Courtesy of Robert Bosch GmbH)

In the above equation, the r_{ii} are the elements of the rotation matrix:

$$\mathbf{R}_{\kappa\varphi\omega} = \begin{bmatrix} \cos(\varphi)\cos(\kappa) & \sin(\omega)\sin(\varphi)\cos(\kappa) + \cos(\omega)\sin(\kappa) & -\cos(\omega)\sin(\varphi)\cos(\kappa) + \sin(\omega)\sin(\kappa) \\ -\cos(\varphi)\sin(\kappa) & -\sin(\omega)\sin(\varphi)\sin(\kappa) + \cos(\omega)\cos(\kappa) & \cos(\omega)\sin(\varphi)\sin(\kappa) + \sin(\omega)\cos(\kappa) \\ \sin(\varphi) & -\sin(\omega)\cos(\varphi) & \cos(\omega)\cos(\varphi) \end{bmatrix} \quad \text{Eq. 2.3}$$

The three-dimensional rotation matrix used in collinearity equation is given in Eq. 2.3. The parameter f is the equivalent focal length, x_b, y_b are the coordinates of the image points, x_h, y_h are the coordinates of the principal point, X_p, Y_p, Z_p and X_o, Y_o, Z_o are the coordinates of object points and projection centre, i.e. camera in the object coordinates system. Parameters d_x and d_y describe the lens distortion model. The camera rotation is described in a rotation matrix with elements $r_{11}..r_{33}$ which correspond to rotation angles $(\omega, \varphi, \kappa)$ and they could also be described as four quaternion rotation elements (q_0, q_1, q_2, q_3) (Hartley & Zisserman, 2011). We may also use simpler mathematical models which describe the 3D-2D relations and use them to initialize the collinearity equation such as 3D homographic transformations or DLT functions (Luhmann, et al., 2011). In order to use LSE systems to solve any of these mathematical models we should linearize the equations by deriving the partial derivations. These mathematical solutions are used in many computer vision and photogrammetry applications in order to estimate the camera position or to reconstruct the three dimensional world using multiple images taken from different views of an object in single or multi-camera systems (e.g. Stereo-Vision systems).

³ Sometimes also called the „collinearity condition equation“

2.4.2. Geometry of Omni-Directional Projection

Omni-Directional objectives or fisheye cameras are widely used in photogrammetry and machine-vision applications such as driver-assistant systems. As already discussed in [2.1], the optical projection of a fisheye camera with an extreme field-of-view is different than a central-projection camera thus creates the necessity of using different projection models to describe the lens geometry.

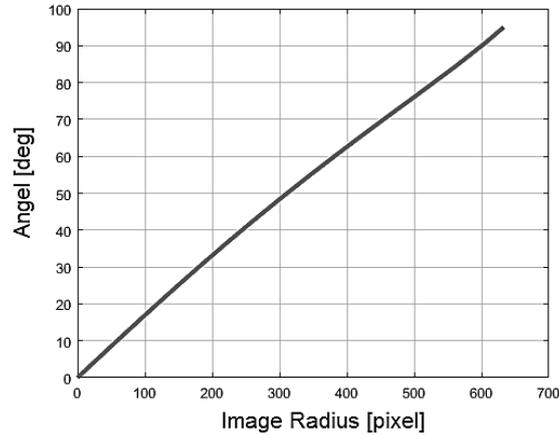


Fig. 2.17. Example of the projection curve of a fisheye camera
(Courtesy of Robert Bosch GmbH)

Most of the classical models used for fisheye calibrations are radial-angular based transformation functions because of the almost radial symmetric projection geometry from which the image coordinates could be derived by considering e.g. the radius of the projected object point using trigonometric functions. In this chapter we study some of the classical and recent models for describing the fisheye optical projection.

2.4.2.1. Equidistant Model:

This model, described as $r' \sim \varphi$, can - to some extents - model the characteristics of fisheye lenses (Schröder, 1981) (Xiong & Turkowski, 1997) (Schneider, et al., 2009). The image coordinates are therefore calculated as follows (Abraham & Förstner, 2005):

$$\begin{aligned} x_b &= c \cdot S_{xy} \cdot \frac{X_p}{R_p} \operatorname{atan}\left(\frac{R_p}{Z_p}\right) + x_h + d_x \\ y_b &= c \cdot \frac{Y_p}{R_p} \operatorname{atan}\left(\frac{R_p}{Z_p}\right) + y_h + d_y, \quad R_p = \sqrt{X_p^2 + Y_p^2} \end{aligned} \quad \text{Eq. 2.4.}$$

In Eq. 2.4, the parameter S_{xy} is the aspect ratio of x and y axis, and is further discussed in chapter 2.4.3. This is a simple model which in some cases may not provide us the necessary calibration accuracies. This means that the actual camera projection behaviour can be different from the equidistant projection model. For example, when using cameras with extreme opening angles and complex radial distortions, we can potentially see remaining errors higher than 20 pixels. This could result in modelling accuracies which are not acceptable for many applications (Abraham, 2012).

2.4.2.2. Stereographic Projection:

This model is based on the $r' \propto \tan(\alpha \cdot \varphi)$ mapping. Therefore the image coordinates would be calculated from object coordinates as follows (Abraham & Förstner, 2005) (Schneider, et al., 2009):

$$\begin{aligned}
 r_b &= c \cdot \tan(\alpha \cdot \varphi) & R_p &= \sqrt{X_p^2 + Y_p^2} \\
 x_b &= c \cdot S_{xy} \cdot \frac{X_p}{R_p} \tan\left(\alpha \cdot \operatorname{atan}\left(\frac{R_p}{Z_p}\right)\right) + x_h + d_x \\
 y_b &= c \cdot \frac{Y_p}{R_p} \tan\left(\alpha \cdot \operatorname{atan}\left(\frac{R_p}{Z_p}\right)\right) + y_h + d_y
 \end{aligned} \tag{Eq. 2.5}$$

This model can describe the projection of fisheye cameras, but may not deliver the necessary accuracies when using cameras with extreme opening angles and projection behaviours different from the tangent function. Therefore we can potentially have errors higher than 20 pixels (Abraham, 2012).

2.4.2.3. Extended Modified Orthogonal Projection:

This orthogonal model is based on $r' \sim \sin(\alpha \cdot \varphi)$ projection (Abraham & Förstner, 2005) and is similar to the equisolid-angle projection (Schneider, et al., 2009) with variable parameters. Therefore the image coordinates is calculated using the following equations:

$$\begin{aligned}
 r_b &= c \cdot \sin(\alpha \cdot \varphi) & R_p &= \sqrt{X_p^2 + Y_p^2} \\
 x_b &= c \cdot S_{xy} \cdot \frac{X_p}{R_p} \sin\left(\alpha \cdot \operatorname{atan}\left(\frac{R_p}{Z_p}\right)\right) + x_h + d_x \\
 y_b &= c \cdot \frac{Y_p}{R_p} \sin\left(\alpha \cdot \operatorname{atan}\left(\frac{R_p}{Z_p}\right)\right) + y_h + d_y
 \end{aligned} \tag{Eq. 2.6}$$

The extension and modification of the orthogonal projection is the combination of different functions which together can describe the projection of omnidirectional lenses:

$$r' = c \cdot \sin(\alpha \cdot \varphi) + B \cdot \varphi^2 + C \cdot \varphi^4 \tag{Eq. 2.7}$$

The accuracy of the extended modified orthogonal projection can reach very good projection-modelling results below 1 pixel depending of the optical mapping of the camera (Schneider, et al., 2009). However the modelling accuracy can in some cases decrease to 5 or 10 pixels if a unique or complex fisheye projection is being modelled (Abraham, 2012). Therefore, this model is suitable for accurate camera calibration based on first studying the optical characteristics of the camera.

2.4.2.4. D. Scaramuzza Omni-Directional Model:

This model, introduced by Davide Scaramuzza, is a calibration model designed for catadioptric⁴ and fisheye cameras. This model is based on the assumption that the optical projection or the imaging function can be described by a Taylor series expansion. (Scaramuzza, et al., 2006) (Scaramuzza, 2014).

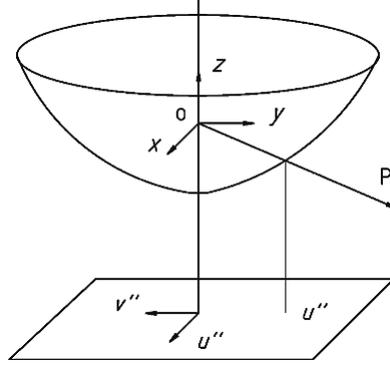


Fig. 2.18. Coordinate system of the Scaramuzza omnidirectional model in the catadioptric case
Illustrated after (Scaramuzza, et al., 2006)

$$\mathbf{u}'' = \mathbf{A}\mathbf{u}' + \mathbf{t} \quad , \quad \mathbf{P} = \mathbf{g}(\mathbf{u}'') = \mathbf{g}(\mathbf{A}\mathbf{u}' + \mathbf{t})$$

Eq. 2.8

$$\mathbf{g}(\mathbf{u}'', \mathbf{v}'') = (\mathbf{u}'', \mathbf{v}'', f(\mathbf{u}'', \mathbf{v}''))$$

In Eq. 2.8, \mathbf{P} is the point in object space and \mathbf{g} is the projection function. The matrices $\mathbf{A}_{2 \times 2}$ and $\mathbf{t}_{2 \times 1}$ are used as an affine transformation to transform the sensor coordinates of $(\mathbf{u}', \mathbf{v}')$ to the camera coordinates $(\mathbf{u}'', \mathbf{v}'')$. Instead of having a specific model for the function f based on the design of lens or mirror, a generalized parametrization is used in the form of a polynomial based on image radius ρ'' : (Scaramuzza, et al., 2006)

$$f(\mathbf{u}'', \mathbf{v}'') = a_0 + a_2 \rho''^2 + \dots + a_N \rho''^N \quad , \quad a_1 = 0$$

Eq. 2.9

The generalization of function f in the Scaramuzza model is given in Eq. 2.9. The coefficients a_i and the polynomial degree N and will be estimated by the calibration. The variable a_1 is set to zero to satisfy the condition $df/d\rho = 0$ when $\rho = 0$. This holds based on studying the function f for hyperbolic and parabolic mirrors and also fisheye cameras (Scaramuzza, et al., 2006). The Scaramuzza model can provide acceptable accuracy for most of the omnidirectional cameras. For example, this model can deliver up to around 0.3 pixels for special catadioptric cameras (Scaramuzza, et al., 2006). In some cases using cameras that have opening angles of around 195 [deg] or higher, and having complex optical projection and lens distortion, the accuracy of the camera calibration could be around 1 to 2 pixels (Abraham, 2012). This accuracy is still good and enough for many applications but is not sufficient to fulfil the required sub-pixel accuracies of the driver-assistant systems and their corresponding applications.

⁴ Catadioptric cameras are made up of a hyperbolic mirror and a perspective camera

2.4.2.5. C. Mei Omni-Directional Model:

Introduced by Christopher Mei, this model is an optimal solution for modelling the projection and distortion of fisheye cameras with practically no limit for opening angles (Mei & Rives, 2007). This model (noted as C. Mei model) provides overall high accuracies compared to other omni-directional camera models. Using this model for calibrating fisheye cameras with for example FOV higher than 195 [deg] and complex radial distortions, one can reach accuracies (RMS) of around 0.2 pixel (Abraham, 2012). Therefore, the C. Mei is a very suitable camera model for applications that benefit from accurate sub-pixel image measurements, for which the high accuracy of calibration is a key factor.

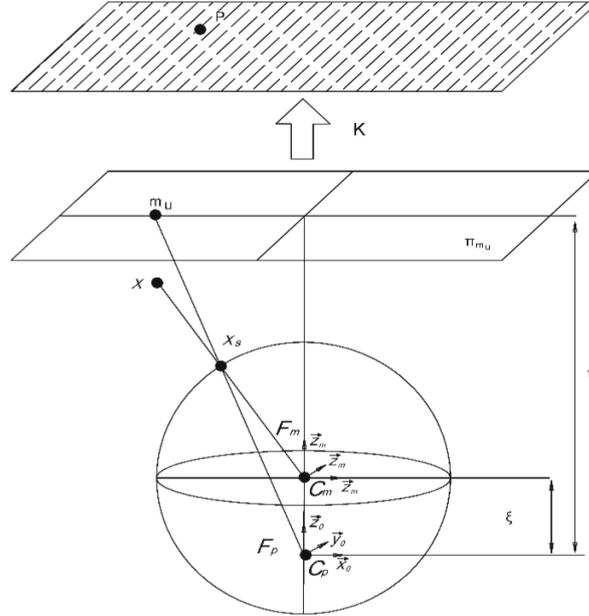


Fig 2.19. The world-to-camera projection of C.MeI model
Illustrated after (Mei & Rives, 2007)

$$P_s = (X_s, Y_s, Z_s) = \frac{(X_p, Y_p, Z_p)}{\|(X_s, Y_s, Z_s)\|} \quad , \quad m = (x_m, y_m) = \left(\frac{X_s}{Z_s + \xi}, \frac{Y_s}{Z_s + \xi}, 1 \right)$$

Eq. 2.10

$$L(\rho) = 1 + k_1\rho^2 + k_2\rho^4 + k_5\rho^6 \quad , \quad \rho = \sqrt{x^2 + y^2}$$

As illustrated in Fig 2.19, in the C. Mei model, the camera projection considers an angular projection on a sphere. The points are projected on a unit sphere and shifted into a reference frame with the centre of C_p . The points are then projected onto a normalized plane (m_u) and the distortion $L(\rho)$ is considered. In Eq. 2.10 the parameters X_p, Y_p and Z_p are the 3D point coordinates, P_s is the unified 3D coordinates, m is the projected image points in normalized plane and $k_1..k_3$ are the radial distortion parameters (Mei & Rives, 2007) and are discussed in chapter [2.4.3]. We will further use this model as the basis model to introduce an extended omnidirectional camera model in chapter 3.

The features and effects of camera optical module or the deviation of those features with respect to the world-to-camera projection could be divided in different optical effects such as distortion, chromatic aberration and so on. The optical distortion is mostly by far the biggest effect here and is needed to be analysed and modelled. This processes of modelling the camera optical distortion and different distortion models are studied in the next chapter.

2.4.3. Calibration Distortion Models

Since the introduction of the Brown calibration model (Brown, 1971) there have been many calibration models introduced for different cameras and lenses with different advantages and disadvantages. In this chapter we discuss some of the current available calibration models which are classically used or recently introduced.

2.4.3.1. Brown distortion model:

The first calibration model acceptably suitable for calibrating the central projection cameras is the Brown-Fraser model (Luhmann, et al., 2011) (Fraser, 1997). This model is called a physical model because it is based on the physical behaviour of camera lenses and attempts to model different existing distortions based on their physical interpretations. The model consists of radial distortion together with a tangential distortion, aspect and skew factors:

Parameters	Meaning/Interpretation
C_y, S_{xy} or C_x, C_y	Equivalent Focal Length, Aspect Ratio= $S_{xy} = \left(\frac{C_x}{C_y}\right)$
x_h, y_h	Principal-Point Coordinates
A_1, A_2, A_3	Radial Distortion
B_1, B_2	Tangential Distortion
C_2	Skew Factor (the affine shear of axes)

Table 2.1. Parameters of Brown distortion model

Considering the focal length parameter, it is beneficial to use two separate parameters for x and y coordinates using the aspect ratio S_{xy} . In the era of analogue photogrammetry this parameter helped to compensate the scale error in the film in x and y direction (Moffitt & Mikhail, 1980) due to tensions and mechanical procedures or the possible scale error during the digitalization process. In digital photogrammetry, this will compensate axial scale difference due to the possible minimal errors in image sensor lithography and the production tolerances of the optical module. After the 3D-to-2D (world-to-camera) projection using C_x and C_y (or respectively: C_y, S_{xy}) we have the undistorted projected points (x_u, y_u) and can apply the distortion using the additional parameters (APs).

By having a look at Eq. 2.2 we can see that the projected image points are in the same unit as the focal length. This means that if the image observations are handled in pixel units, the focal length should be also in pixel units and could be converted to and from [mm] by having the pixel size. Furthermore, using the value 1 for the focal length (C_y) in the initial projection, one can also alternatively perform the distortion correction with the normalized coordinates, and then use the equivalent focal length values later together with skew correction to calculate the distorted image coordinates x_d and y_d from the normal plane.

The Brown distortion model is as follows. First, the radial distortion parameters are considered:

$$\begin{aligned} dx_{ru} &= x_u \cdot (A_1\rho^2 + A_2\rho^4 + A_3\rho^6) \\ dy_{ru} &= y_u \cdot (A_1\rho^2 + A_2\rho^4 + A_3\rho^6) \\ \rho &= \sqrt{x_u^2 + y_u^2} \end{aligned} \tag{Eq. 2.11}$$

Then, parallel to the radial distortion correction we correct the tangential distortion:

$$\begin{aligned} dx_{tu} &= 2B_2x_u y_u + B_1(\rho^2 + 2x_u^2) \\ dy_{tu} &= 2B_1x_u y_u + B_2(\rho^2 + 2y_u^2) \end{aligned} \tag{Eq. 2.12}$$

Now we correct the coordinates using the radial and tangential distortion values:

$$x_{d_{rt}} = x_u + dx_{ru} + dx_{tu} \qquad y_{d_{rt}} = y_u + dy_{ru} + dy_{tu} \tag{Eq. 2.13}$$

Finally, after adding the calculated distortion values to the projected coordinates, the principal point and the skew parameters are used. In case of using normal projected coordinates, the equivalent focal length is used here to scale and map the points to pixel coordinates:

$$\begin{aligned} \begin{bmatrix} x_d \\ y_d \end{bmatrix} &= \begin{bmatrix} x_{d_{rt}} + C_2 y_{d_{rt}} + x_h \\ y_{d_{rt}} + y_h \end{bmatrix} & \begin{bmatrix} x_d \\ y_d \end{bmatrix} &= \begin{bmatrix} C_x(x_{d_{rt}} + C_2 y_{d_{rt}}) + x_h \\ C_y y_{d_{rt}} + y_h \end{bmatrix} \end{aligned} \tag{Eq. 2.14}$$

In Eq. 2.14, the corrections are given with the skew parameter (left) or with the skew and normal-to-pixel scaling using the focal length (right). Based on Eq. 2.11 and Eq. 2.12 we can reformulate the distortion equation:

$$\begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} = \begin{bmatrix} A_1\rho^2 + A_2\rho^4 + A_3\rho^6 + 2B_2y_u & 0 & B_1(\rho^2 + 2x_u^2) \\ 0 & A_1\rho^2 + A_2\rho^4 + A_3\rho^6 + 2B_1x_u & B_2(\rho^2 + 2y_u^2) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} \tag{Eq. 2.15}$$

The representation of radial and tangential distortion using matrix multiplications is given in Eq. 2.15 which could be used as a more compact representation.

2.4.3.2. Brown-Fryer distortion model:

It is possible to use a balanced radial distortion in Brown model (Clarke & Fryer, 1998) (Luhmann, et al., 2011) to model the camera radial distortion, which is sometimes called Brown-Fryer model:

$$\begin{aligned} dx_{ru} &= x_u \cdot (A_1(\rho^2 - \rho_0^2) + A_2(\rho^4 - \rho_0^2) + A_3(\rho^4 - \rho_0^6)) \\ dy_{ru} &= y_u \cdot (A_1(\rho^2 - \rho_0^2) + A_2(\rho^4 - \rho_0^2) + A_3(\rho^4 - \rho_0^6)) \end{aligned} \quad \text{Eq. 2.16}$$

The balanced radial distortion could be advantageous for real-time systems by potentially reducing the necessary amount of device memory used for allocation of distortion data, but does not significantly affect the calibration accuracy of digital cameras. This is because the absolute amount of the modelled distortion is smaller while considering the estimated equivalent focal length (Luhmann, et al., 2011). Furthermore, if the radial distortion curve of a model becomes un-invertible, this model enables us - to some extent - to manually select the parameter ρ_0 , thus changing the distortion curve's infimum and affect the un-invertibility in a defined domain. If the distortion function becomes not invertible in the interested image area, the use of ρ_0 is not preferable in cases where undistortion needs to be calculated from the distortion function. The next models are designed to further model the local distortions and image deformations by using extra parameters in the distortion model.

2.4.3.3. Brown Mixed-AP distortion model:

As an extension to Brown model, this model is a combination of mathematical polynomials and physical parameters, which increases the flexibility of an only-physical model to fit itself to different image distortions and deformations. The total parameter-count of this model is 21 (Brown, 1976).

$$\begin{aligned} dx_u &= a_1x + a_2y + a_3xy + a_4y^2 + a_5x^2y + a_6xy^2 + a_7x^2y^2 \\ &\quad + \frac{x}{c} [a_{13}(x^2 - y^2) + a_{14}(x^2y^2) + a_{15}(x^4 - y^4)] \\ &\quad + x[a_{16}(x^2 + y^2) + a_{17}(x^2 + y^2)^2 \\ &\quad + a_{18}(x^2 + y^2)^3] + a_{19} + a_{21} \left(\frac{x}{c} \right) \end{aligned} \quad \text{Eq. 2.17}$$

$$\begin{aligned} dy_u &= a_8xy + a_9x^2 + a_{10}x^2y + a_{11}xy^2 + a_{12}x^2y^2 \\ &\quad + \frac{y}{c} [a_{13}(x^2 - y^2) + a_{14}(x^2y^2) + a_{15}(x^4 - y^4)] \\ &\quad + y[a_{16}(x^2 + y^2) + a_{17}(x^2 + y^2)^2 \\ &\quad + a_{18}(x^2 + y^2)^3] + a_{20} + a_{21} \left(\frac{y}{c} \right) \end{aligned}$$

In Eq. 2.17 the parameters $a_1..a_7$ and $a_8..a_{12}$ are the in-plane distortion components in x and y directions, $a_{13}..a_{15}$ are the out-of-plane distortion components, $a_{16}..a_{18}$ are radial distortion parameters, and $a_{19}..a_{21}$ are the interior orientation correction for the principal point and the focal length.

2.4.3.4. Ebner distortion model:

In 1976 Ebner introduced a model which could fit to certain image local distortions and describe image deformations based on bivariate second-order orthogonal polynomials with 12 parameters. This model is mostly suitable for aerial applications - when for multiple images the image base and the forward overlap is known - especially when the measurement points cover almost the entire image (Ebner, 1976) (Clarke & Fryer, 1998).

$$\begin{aligned} dx_u &= xb_1 + yb_2 - 2kb_3 + xyb_4 + lb_5 + xlb_7 + ykb_9 + klb_{11} \\ dy_u &= -yb_1 + xb_2 + xyb_3 - 2lb_4 + kb_6 + ykb_8 + ylb_{10} + klb_{12} \end{aligned} \quad \text{Eq. 2.18}$$

$$k = x^2 - \frac{2}{3}b^2 \quad , \quad l = y^2 - \frac{2}{3}b^2$$

In Eq. 2.18, the parameter b is chosen correspondingly to the base distance in image space. This parameter is dependent on image dimensions and – in case of aerial images – the forward overlap.

2.4.3.5. Grün distortion model:

The model from Ebner was further extended in 1978 by Grün to a 4th order orthogonal polynomial which can relatively model complex image deformations (Grün, 1978).

$$\begin{aligned} dx_u &= xa_1 + ya_2 + xya_3 + la_4 - \frac{10}{7}ka_6 + xpa_7 + yka_8 + xla_9 + yqa_{10} + ra_{15} + xypa_{16} \\ &\quad + kla_{17} + xyqa_{18} + sa_{19} + yra_{25} + xlp a_{26} + ykqa_{27} + xsa_{28} + lra_{33} \\ &\quad + xypqa_{34} + ksa_{35} + yqra_{39} + xpsa_{40} + rsa_{43} \end{aligned} \quad \text{Eq. 2.19}$$

$$\begin{aligned} dy_u &= -ya_1 + xa_2 - \frac{10}{7}la_3 + ka_5 - xya_6 + xpa_{11} + yka_{12} + xla_{13} + yqa_{14} + ra_{20} \\ &\quad + xypa_{21} + kla_{22} + xyqa_{23} + sa_{24} + yra_{29} + xlp a_{30} + ykqa_{31} + xsa_{32} \\ &\quad + lra_{36} + xypqa_{37} + ksa_{38} + yqra_{41} + xpsa_{42} + rsa_{44} \end{aligned}$$

The parameters k, l, p, q, r and s are defined as:

$$\begin{aligned} k &= x^2 - \frac{1}{2}b^2 \quad , \quad l = y^2 - \frac{1}{2}b^2 \\ p &= x^2 - \frac{17}{20}b^2 \quad , \quad q = y^2 - \frac{17}{20}b^2 \\ r &= x^2 \left(x^2 - \frac{31}{28}b^2 \right) + \frac{9}{70}b^4 \quad , \quad s = y^2 \left(y^2 - \frac{31}{28}b^2 \right) + \frac{9}{70}b^4 \end{aligned} \quad \text{Eq. 2.20}$$

Concerns have been raised for an overestimation possibility, especially the correlation between the extrinsic and intrinsic calibration parameters. Furthermore mathematical models are sometimes not the favourable solution as they have little physical interpretation when modelling the image distortion. Thus the use of a physically-motivated model is sometimes preferred (Tang, et al., 2012 B).

2.4.3.6. Tschebycheff distortion model:

Another model often used for camera calibration is the Tschebycheff calibration model (Abraham, 2004) (Mason & Handscomb, 2002). We can use the Tschebycheff polynomials⁵ of the second or third order to model the lens distortions. These orthogonal polynomials can model an anonymous function with normal coordinates/values without having any knowledge of the physical properties:

$$\begin{aligned} dx &= \sum_r \sum_c a_{rc} P_c(x) P_r(y) \\ dy &= \sum_r \sum_c b_{rc} P_c(x) P_r(y) \end{aligned} \quad \text{Eq. 2.21}$$

Therefore the model with the scaled/normalized coordinates looks like:

$$\begin{aligned} dx &= \sum_r \sum_c a_{rc} T_c(T_{sx} \cdot x_b) T_r(T_{sy} \cdot y_b) \\ dy &= \sum_r \sum_c b_{rc} T_c(T_{sx} \cdot x_b) T_r(T_{sy} \cdot y_b) \end{aligned} \quad \text{Eq. 2.22}$$

$$T_{sx} = \frac{2}{\text{Image Width}} \quad , \quad T_{sy} = \frac{2}{\text{Image Height}}$$

In order to calculate the Tschebycheff polynomials we can use their recurrence relation characteristic. By using the following relations, we can generate all the Tschebycheff polynomials in a recursive fashion (Mason & Handscomb, 2002) (Scheu, 1995):

$$\begin{aligned} T_0(x) &= 1 \quad , \quad T_1(x) = x \\ T_{n+1}(x) &= 2x T_n(x) - T_{n-1}(x) \end{aligned}$$

$$a = \begin{bmatrix} 0 & P_0 & P_4 & P_{11} \\ 0 & P_3 & P_9 & 0 \\ P_1 & P_7 & 0 & 0 \\ P_5 & 0 & 0 & 0 \end{bmatrix}_{N+1, N+1} \quad , \quad b = \begin{bmatrix} 0 & 0 & -P_3 & P_{12} \\ P_0 & -P_1 & P_{10} & 0 \\ P_2 & P_8 & 0 & 0 \\ P_6 & 0 & 0 & 0 \end{bmatrix}_{N+1, N+1} \quad \text{Eq. 2.23}$$

In Eq. 2.23 the Tschebycheff model polynomials of order 3 is given. Two parameters in each matrix are set to zero because it would otherwise create a direct correlation between the corresponding above coefficients. The intrinsic model would finally look like:

$$C_y, S_{xy}, x_h, y_h, T_{sx}, T_{sy}, Tk_6, Tk_7, \dots, Tk_{18} \quad , \quad P_i = Tk_{i+6} \quad \text{Eq. 2.24}$$

The C_y, S_{xy} represent the equivalent focal length and aspect ratio, and x_h, y_h represent the coordinates of the principal point in x and y directions, and the $P_i = Tk_{i+6}$ parameters are the estimated factors of the Tschebycheff polynomials.

⁵ Named after Pafnuty Lvovich Tschebychev, mathematician, inventor of several mathematical and statistical theorems

2.4.3.7. Zernike distortion model:

As another calibration model, we can further name the Zernike model. This model is based on the Zernike polynomials⁶ that are orthogonal complex polynomials defined on a plane on the area of a unit circle or unit disk ($0 \leq r \leq 1$). Therefore, polar coordinates are used in these functions (Born & Wolf, 1999) (Abraham, 2004):

$$V_n^m(r, \varphi) = R_n^m(r) \cdot e^{jm\varphi} \quad \text{Eq. 2.25}$$

The R_n^m in Eq. 2.25 consists of a set of orthogonal polynomials as functions of r . We can further divide the complex part in to two parts, one imaginary and one real part ($e^{jm\varphi} = \cos(m\varphi) + j \sin(m\varphi)$) and the R_n^m could also be divided in two odd and even set. Therefore we would finally have:

$$\begin{aligned} {}^oV_n^m(r, \varphi) &= R_n^m(r) \cdot \sin(m\varphi) \\ {}^eV_n^m(r, \varphi) &= R_n^m(r) \cdot \cos(m\varphi) \end{aligned} \quad \text{Eq. 2.26}$$

Dividing of Zernike polynomials into real and complex sets is given in Eq. 2.26. The equations for both odd and even sets are later derived (Born & Wolf, 1999). Using these functions, we can write the Zernike calibration equations. The parameters $a_{00}V_0^0$, $a_{11o}V_1^1$ and $a_{11e}V_1^0$ have to be avoided since they have direct correlation with other intrinsic/extrinsic calibration parameters.

$$\begin{aligned} dx &= (x - x_0) \sum_{n,m} a_{nm} V_n^m(x - x_0, y - y_0) \\ dy &= (y - y_0) \sum_{n,m} b_{nm} V_n^m(x - x_0, y - y_0) \end{aligned} \quad \text{Eq. 2.27}$$

Eq. 2.27 gives the description of Zernike model based on image coordinates. Because the derived Zernike equations are functions of trigonometric functions (Born & Wolf, 1999), they would not be suitable for real-time applications, especially for mobile applications and compact machine-vision systems. This is because the estimation of all of those trigonometric functions and the Zernike functions in real-time (in milliseconds) needs devices with high processing power which are mainly not available in those systems which counts as a practical limitation for this model. Furthermore, the Zernike model is a mathematical model and there is basically very little physical interpretation of the image deformations and distortions estimated by the parameters of this model.

⁶ Named after Fritz Zernike, optical physicist, winner of the 1953 physics Nobel Prize

2.4.3.8. Legendre distortion model:

As recently introduced (Tang, et al., 2012 A), this model is based on the Legendre orthogonal polynomials⁷ and thus called Legendre APs. The L_i functions are orthogonal Legendre polynomials defined on normalized image coordinates (Tang, et al., 2012 A). Therefore:

$$l_m(x, b_x) = L_m\left(\frac{x}{b_x}\right), \quad l_n(y, b_y) = L_n\left(\frac{y}{b_y}\right) \quad \text{Eq. 2.28}$$

$$L_0(x) = 1, \quad L_1(x) = x, \quad L_2(x) = \frac{1}{2}(3x^2 - 1), \quad L_3(x) = \frac{1}{2}(5x^3 - 3x), \dots$$

The parameter b_x in Eq. 2.28 is half of the respective image dimension. The two dimensional bivariate OPs ($P_{m,n}$) are the multiplication of two l_i functions of x and y with different degrees:

$$P_{m,n} = S \cdot f_{m,n} \quad \text{Eq. 2.29}$$

$$f_{m,n} = f_{m,n}(x, y, b_x, b_y) = l_m(x, b_x) l_n(y, b_y)$$

The scale factor S scales the relatively small distortion values, which are usually not more than few pixels, for numerical stability. Based on the order of the Legendre APs distortion model, the $P_{m,n}$ could be selected in lexicographical order:

$$\begin{aligned} & P_{00} \\ & P_{10}, P_{01} \\ & P_{20}, P_{11}, P_{02} \\ & P_{30}, P_{21}, P_{12}, P_{03} \\ & P_{40}, P_{31}, P_{22}, P_{13}, P_{04} \\ & \dots \end{aligned} \quad \text{Eq. 2.30}$$

For example, when using Legendre APs for $M_x = M_y = 4$ and $N_x = N_y = 3$ we would have a total of 34 parameters ($a_1 \dots a_{34}$):

$$\begin{aligned} \Delta x &= a_1 P_{10} + a_2 P_{01} + a_3 P_{20} + a_4 P_{11} + a_5 P_{02} + a_6 P_{30} + a_7 P_{21} + \dots \\ \Delta y &= a_2 P_{10} - a_1 P_{01} + a_{20} P_{20} - a_3 P_{11} - a_4 P_{02} + a_{21} P_{30} + a_{22} P_{21} + \dots \end{aligned} \quad \text{Eq. 2.31}$$

The calculation of the distortion, based on the Legendre coefficients is given in Eq. 2.31. We should mention and consider the high correlation between $P_{10}, P_{01}, P_{20}, P_{11}$ and P_{02} in $\Delta x, \Delta y$ when performing model parameterization. The final accuracy of Legendre APs model, though showing less correlation between calibration parameters, is comparable to the Brown Mixed AP and Grün using DMC camera for airborne images (Tang, et al., 2012 A).

⁷ Named after mathematician Adrien-Marie Legendre.

2.4.3.9. Fourier distortion model:

This recently introduced model is based on Fourier series to model the image distortion (Tang, et al., 2012 B). The bivariate Fourier series are as follows:

$$c_{mn} = S \cdot \cos(\mu \pm nv), \quad s_{mn} = S \cdot \sin(\mu \pm nv)$$

$$@ \quad u = \pi \frac{x}{b_x}, \quad v = \pi \frac{y}{b_y}, \quad S = 10^{-6} \quad (\text{example for normalization factor}) \quad \text{Eq. 2.32}$$

$$\cos(\mu \pm nv), \sin(\mu \pm nv) \quad @ \quad m, n = 0, 1, 2, 3, \dots, \quad u, v \in [-\pi, \pi]$$

Description of the Fourier model based on the Fourier series is given in Eq. 2.32 where b_x and b_y are half of the respective image dimension and S is the scale factor for numerical stability. The Fourier APs would form as follows:

$$\Delta x = \sum_{m=1}^M \sum_{n=-N}^N a_{mn} c_{mn} + b_{mn} s_{mn} + \sum_{n=1}^N a_{0n} c_{0n} + b_{0n} s_{0n}$$

$$\Delta y = \sum_{m=1}^M \sum_{n=-N}^N a'_{mn} c_{mn} + b'_{mn} s_{mn} + \sum_{n=1}^N a'_{0n} c_{0n} + b'_{0n} s_{0n}$$

Eq. 2.33

Eq. 2.33 describes the coefficients of the Tschebycheff model, where a_{mn} , b_{mn} , a'_{mn} and b'_{mn} are unknown coefficients that need to be estimated. As an example, the Fourier APs for $M = N = 1$ would look like (16 parameters):

$$\Delta x = a_1 c_{10} + a_2 c_{01} + a_3 c_{1,-1} + a_4 c_{11} + a_5 s_{10} + a_6 s_{01} + a_7 s_{1,-1} + a_8 s_{11}$$

$$\Delta y = a_9 c_{10} + a_{10} c_{01} + a_{11} c_{1,-1} + a_{12} c_{11} + a_{13} s_{10} + a_{14} s_{01} + a_{15} s_{1,-1} + a_{16} s_{11}$$

Eq. 2.34

Eq. 2.34 gives the description of calculation of the distortion using the Fourier coefficients of order 1 for both M and N . Generally, the Fourier APs are mathematically rigorous which make them comparable to physical models. This model has certain advantages such as a high distortion modelling accuracy, high parameter significance and the low correlation between extrinsic and intrinsic orientation relative to the classical models (Tang, et al., 2012 B). However, because of having high parameter count (e.g. 16 or 48) all based on trigonometric functions each being again calculated as a series based on Maclaurin expansion (Kreyszig, 2010), the model is very calculation-intensive and is not most suitable for real-time calculation and distortion correction on mobile devices with few processing resources.

2.4.3.10. Considerations on the distortion models

In the previous chapters we have reviewed several calibration distortion models. Here we need to discuss an important issue on the input data, model degree and modelling accuracy that needs to be addressed when using a polynomial-based distortion model.

Generally, to achieve a high calibration accuracy overall the image up to the image edges, the degree of the distortion model or polynomial could be increased. This increment however could not be arbitrarily very high without fulfilling some requirements on the input information, i.e. image observations. Based on the number and distribution of the observations in the image, this could potentially introduce some difficulties in the distortion estimation. One of the most important criteria in this matter is the Runge's Phenomenon (Dahlquist & Björck, 2003). This phenomenon has been discussed in numerical analysis and refers to the increase in errors at the edges of a domain when trying to fit a high-degree polynomial to a function over that domain. Runge's phenomenon is mostly present in conditions in which the high degree of polynomial comes close to the number of equidistant sampled points in the domain. Because the modelling of lens distortion could be considered as a mathematical function fitting or polynomial fitting (Tang, et al., 2012 A) (Tang, et al., 2012 B) we need to make sure that this phenomenon does not happen when performing the camera calibration.

Therefore, in order to avoid the above issues, we need to increase the number of observations and improve their distribution in the image (up to the image edge), and avoid unnecessary high model or polynomial degrees to make sure the Runge's phenomenon will not happen. The analysis of point distribution on different calibration techniques is performed in [6.2.1] and [6.4.2].

Now we have discussed several calibration distortion models ranging from the classical models to the ones recently introduced. Later in [7.1.1] and [7.2.1] these models will be used in our camera calibration tests and their accuracies will be compared. We now need to study different approaches and techniques used for camera calibration. They consist of conventional solutions such as collimator calibration, bundle-block calibration, test stand calibration and some other methods. These solutions will be discussed separately in the following chapter.

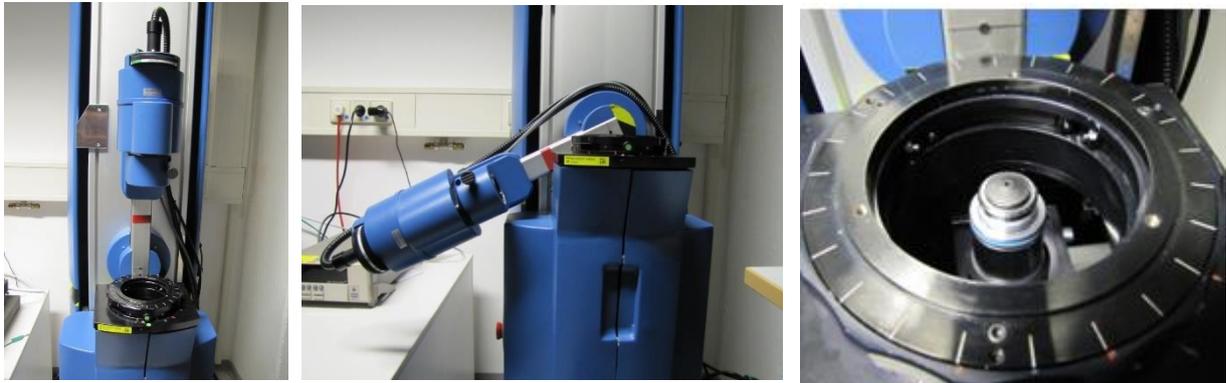
2.5. Conventional Calibration Solutions and Methods

The camera calibration methods could be divided in different major categories regardless of the type of camera objective, which is e.g. central-projection or omni-directional camera. Each category has its own unique features, specifications, and methods for carrying out camera calibration.

In this chapter we briefly study and discuss some of the main conventional calibration methods that are nowadays used in the field of photogrammetry and computer vision. First we study the use of collimators as a classic approach to camera calibration. After that we discuss bundle-block calibration as a standard and widespread solution today. Next we would study the test stand calibration as a fast and single-shot solution for camera calibration and finally we discuss other recent techniques for calibration which may generally not fit in the above categories.

2.5.1. Collimator Calibration (Light-Ray Measurement)

Since the era of analogue aerial photogrammetry, a classical solution for camera calibration has been the use of collimators which is a standard way of calibrating different types of cameras (Moffitt & Mikhail, 1980). It should be mentioned that using this solution, other camera optical parameters such as lens sharpness (e.g. MTF) can also be measured which is advantageous.



a	b	c
---	---	---

 Fig. 2.20. Camera calibration using collimators

a, b: View of a collimator used for camera calibration (Trioptics GmbH)

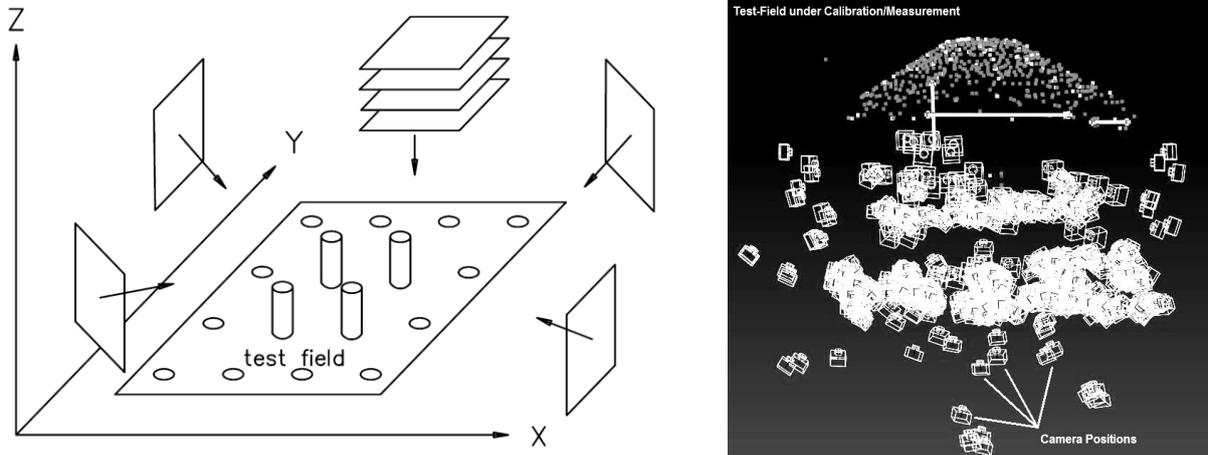
c: View of the objective fixture in the calibration collimator

(Courtesy of Robert Bosch GmbH)

A special advantage of the collimator calibration, illustrated in Fig. 2.20, is that the camera lenses could also be measured separately using a reference image sensor in the collimator, which delivers relatively accurate results for radial lens distortion. The drawback of this solution other than the limitation of objective size is the speed of measurement procedure, which usually takes at least few minutes. It makes this method suitable for calibrating single cameras but not a big number of cameras in a short time.

2.5.2. Bundle-Block Calibration

“The bundle triangulation, also called bundle block adjustment, multi-image triangulation, multi-image orientation is a method for simultaneous numerical fit of an unlimited number of spatially distributed images or bundle of rays” (Luhmann, et al., 2011).



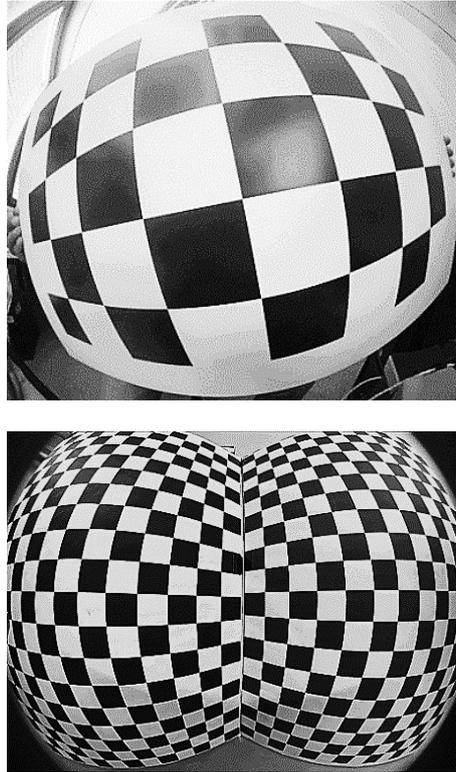
a	b
---	---

 Fig. 2.21. Illustration of close-range bundle-block calibration
a: Illustration of image configuration for test field calibration, Illustrated after (Luhmann, et al., 2011) p. 450
b: Image configurations for bundle-block calibration/measurement of a test-field. Each box is representing one camera position. (Society of measurement technique, VMT GmbH, Courtesy of Robert Bosch GmbH)

Since the introduction of the bundle-block calibration, it has turned into a standard self-calibration solution for many video applications and camera types. Using the sets of images taken from different positions/orientations in the bundle-block adjustment, the three dimensional world could be reconstructed. Bundle-block solution is widely used in today’s applications for camera calibration in close-range and also aerial photogrammetry. As illustrated in in Fig. 2.21 in close-range photogrammetry, it is a very common approach in bundle-block calibration to use test-fields.

The general problem of the correlation between the intrinsic and extrinsic calibration is to overcome by considering a set of defined configuration of camera positions relative to the test-field which introduce geometrical and numerical stability into our bundle-block optimization. In other words, in bundle-block calibration there is the possibility of having regular point distribution in image and strong network-geometry (image constellation) which makes it possible to estimate the intrinsic calibration parameters with high significance and separate them from the image extrinsic orientation parameters.

The strong geometry of bundle-block calibration, when performed with control points in the scene, makes it possible to consider camera calibration models in our calculations as unknown parameters and estimate the calibration parameters when solving the equation system i.e. self-calibration. Self-calibration using bundle-block adjustment has been strongly studied in recent years by introducing new calibration models (Tang, et al., 2012 A) (Tang, et al., 2012 B) which enables the bundle-block calibration to be performed with much improved accuracy and quality.



a Fig. 2.22. Illustration of close-range bundle-block calibration of a fisheye camera using checker-board pattern
b a: Handheld checkerboard-pattern, changing the target position with the camera in a fixed position
 b: Stationary checkerboard pattern, changing the camera position with the target in a fixed position
 (Courtesy of Robert Bosch GmbH)

A special advantage of bundle-block calibration is the ease of use in close-range as changing the position of the camera with respect to the test-field could be relatively trivial. As an example of using close-range bundle-block calibration for a fisheye camera illustrated in Fig. 2.22, we can name the methods using standard checkerboard patterns for which are also many tools available (Mei & Rives, 2007) (Scaramuzza, et al., 2006).

We should also mention the special test-fields with known/measured point coordinates together with using coded-points for automatic camera calibration. As illustrated in Fig. 2.23 these test-fields may also incorporate points located on different planes with different pivot vector angle with respect to each other integrated as a single test-field (Abraham, 2004) (Abraham & Förstner, 2005). This test-field simulates having several captured frames with different orientation angles with respect to the camera. Using the latter has the advantage that fewer image frames need to be captured.

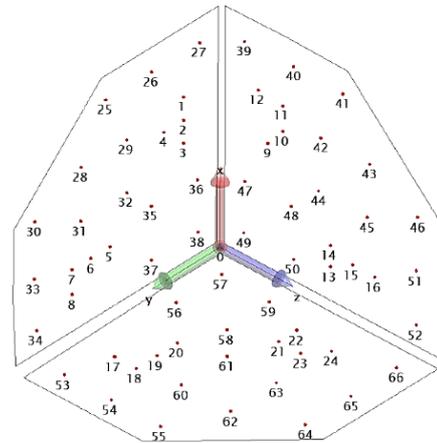
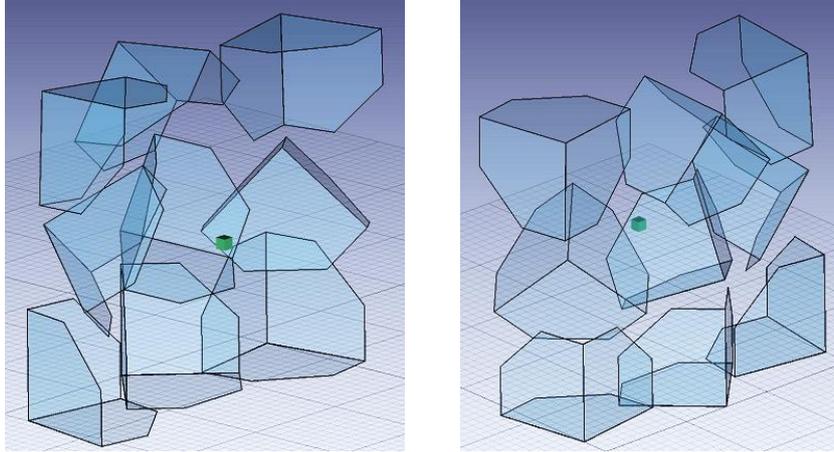
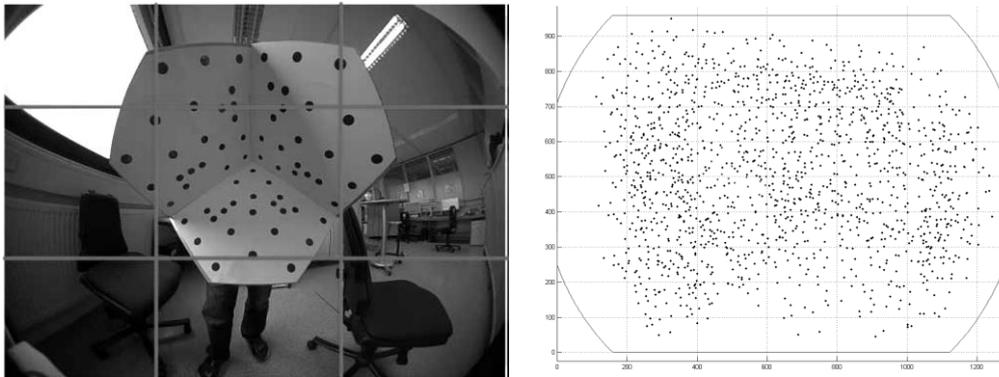


Fig. 2.23. Test-Field with 3 perpendicular planes for fisheye camera calibration



a **b** Fig. 2.24 The example of a simple configuration of the relative position between a fixed fisheye camera (green) and a moving test-field / target (blue) in bundle-block calibration. These positions and the possible positions and orientation between them are used for calibration. a: view from front side, b: view from behind. (Courtesy of Robert Bosch GmbH)

As discussed, the bundle-block calibration can be performed using different configuration of camera extrinsic orientation and is shown in Fig. 2.24. In this case, instead of moving the camera and having the test-field in a fixed position, we move the test-field around the fixed camera. The changes in the relative camera-to-test-field orientation will result in the bundle block calibration benefiting from a very desirable uniform point distribution effect which will especially help us improve the accuracy of calibration and is discussed in 6.2.1. As shown in Fig. 2.25 we can see one captured position of a moving test-field. Images would be captured with the test-field changing its position and orientation continuously.



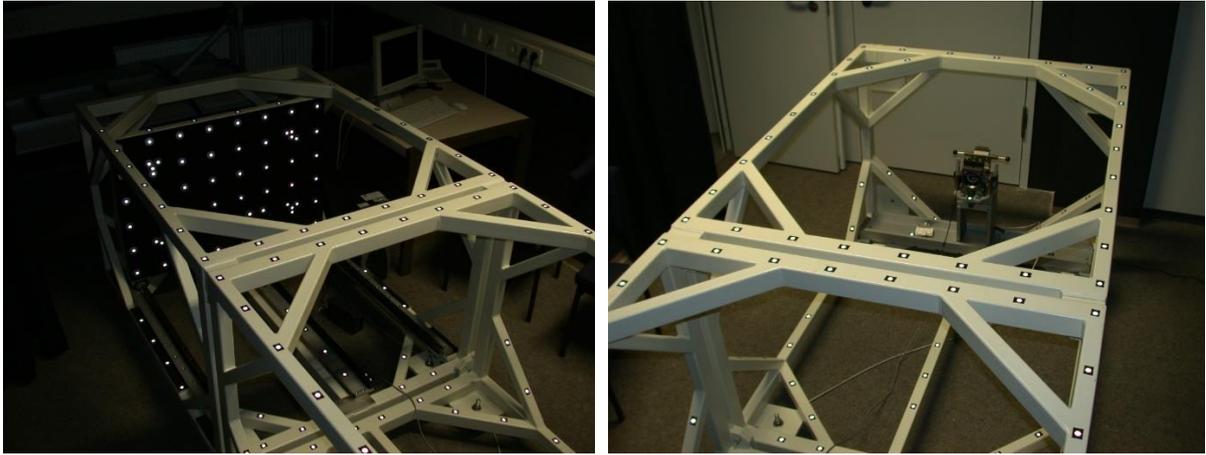
a **b** Fig. 2.25. Illustration of close-range bundle-block calibration of a fisheye camera
a: Example of a captured frame for bundle-block test-field calibration
b: Example of final point distribution after performing the complete bundle-block calibration using multiple images
(Courtesy of Robert Bosch GmbH)

The slight drawback of the bundle-block calibration is that it naturally takes a certain amount of time in order to change the relative camera-to-target position and orientation and capture images from different viewing angles. Therefore this solution is usually not applicable or practical for applications in which the whole calibration process should be performed in few seconds and in short time intervals (e.g. in a camera production line).

2.5.3. Test Stand Single-Shot Calibration

The technique of performing camera calibration using a single image is often called single-shot or test-stand camera calibration. The mathematical principle is the same as bundle-block calibration. It is clear that by having a single image, and having only control points in the image, the complex system of bundle-block equations would be simplified into a spatial resection model with which the additional self-calibration parameters are used.

The hardware platform that is used for this purpose is called a calibration test stand, which has usually four main parts: the test stand frame/structure, the camera mounting fixture, the test-field or target-points with known three dimensional coordinates and the illumination device. Normally in case of using passive illuminated points as target points i.e. retro-reflectors, the illuminating device is installed separately, usually in form of LEDs positioned close to the camera. In case of using active illuminated target points i.e. self-illuminating points, the illuminating device would be integrated in the target point structure.



a	b
---	---

 Fig. 2.26. Test-stand for calibrating central-projection/perspective camera
 a: View of a calibration test stand frame and the target points (test-field)
 b: View of a calibration test stand frame, the camera fixture and the LEDs
 (Courtesy of Robert Bosch GmbH)

In the test stand illustrated in Fig. 2.26 the frame is visible which is marked with retro-reflective points which are used as tie points for photogrammetric calibration of the test stand itself. The test-field with known 3D coordinates is usually located at one end, facing the camera and the illuminating device (LEDs). The LEDs are usually chromatically calibrated with known spectral frequencies, which are compatible with the type of camera/sensor that is mounted in the test stand. Usually the test-field points form a general three-dimensional space, implying that the target points do not lie on one plane. This help decrease intrinsic and extrinsic correlation and increase the significance of the calculated intrinsic parameters.

It is possible to use special point-groups in the target points with unique geometrical topology as shown in Fig. 2.27. This special geometrical design ensures minimal similarity correlation between different point-groups when considering their coordinates on a plane or in image.

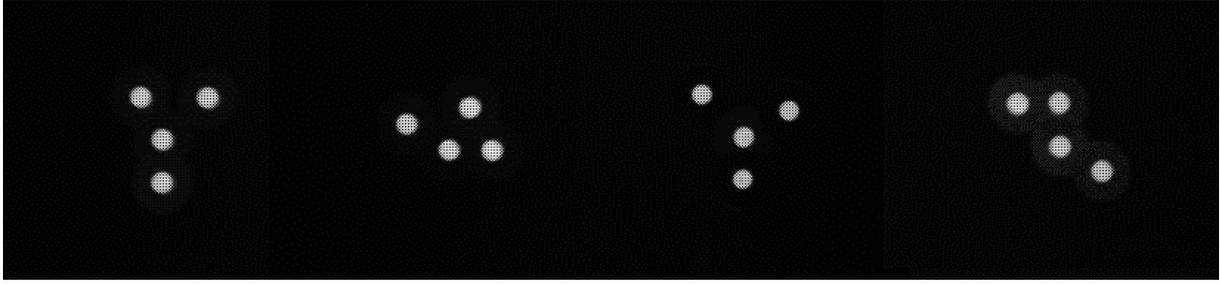


Fig. 2.27. Example of a set of different point-groups using 4 points and different unique geometry (Courtesy of Robert Bosch GmbH) Further discussions on marker measurement in chapter 2.6.

Using the point-groups, a direct correlation analysis of measured points and projected control points makes this possible to match the point-groups between image-space and object-space. The highest correlation between two point-groups or the lowest distance or deviation between two groups during two-dimensional correlation indicates the matching of those points. (Abraham & Hau, 1997). This can be performed automatically and helps the realization of an automated calibration process.

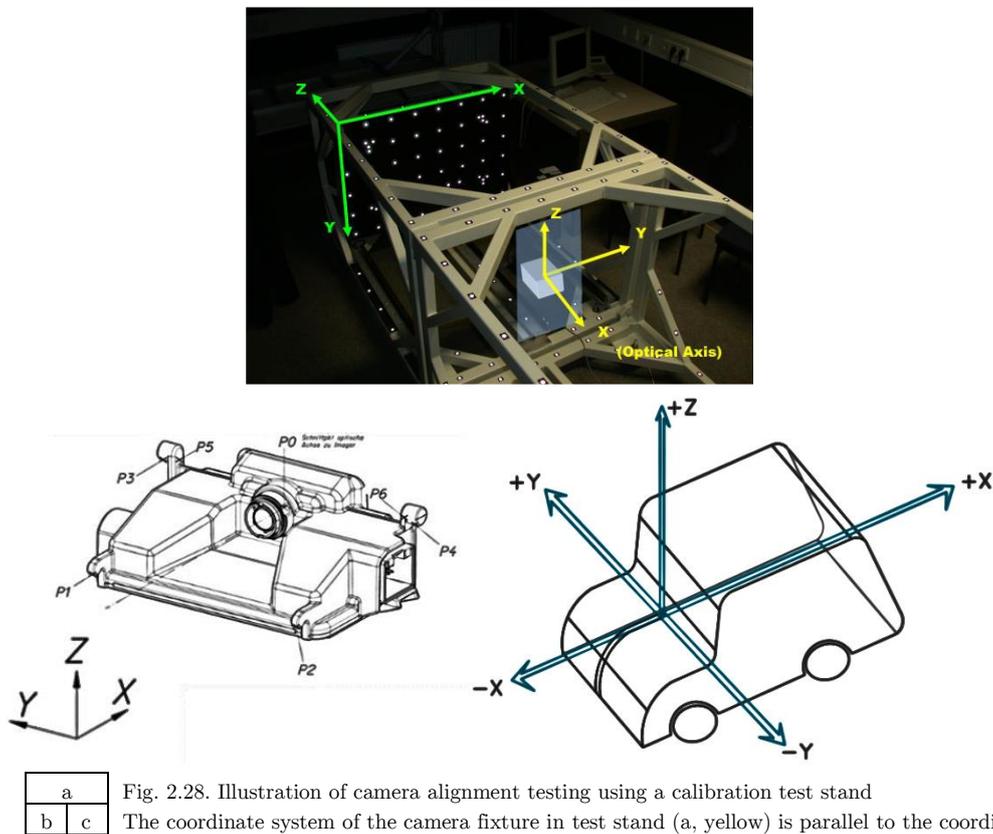


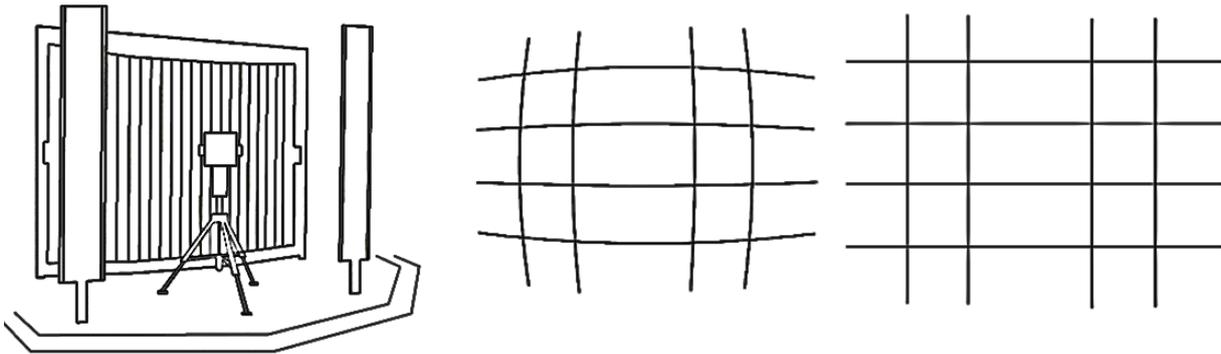
Fig. 2.28. Illustration of camera alignment testing using a calibration test stand

The coordinate system of the camera fixture in test stand (a, yellow) is parallel to the coordinate system of the vehicle (c) illustrated after ISO 4130-1978 (EUR-Lex, 1978) (Courtesy of Robert Bosch GmbH)

Having the 3D coordinates of the test-field point, by capturing a single frame from the camera, one can solve the spatial resection equation system with additional parameters for camera calibration. A special advantage of using the test stand calibration is the possibility of performing camera alignment testing as illustrated in Fig. 2.28. This is realized by calibrating the test stand in a unified process which implies estimating the test-field coordinates in a defined reference camera coordinate system which enables us to estimate the alignment of the imaging components with respect to the reference camera and vehicle axes. A feature that is very hard to achieve in a general bundle-block calibration without camera fixture.

2.5.4. Other Calibration Solutions

One of the classical approaches to camera calibration is plumb-line calibration (Fryer & Brown, 1986). This approach, similar to the background principles of using straight lines in photogrammetry in applications such as “*Planimetry from obliques using perspective grid*” (Moffitt & Mikhail, 1980) considers the projection behaviour of straight lines in the image. In order to model the optical distortion with the most focus on the radial distortion, the deviation of known linear features in object space from a straight line in image space is measured.



a	b	c
---	---	---

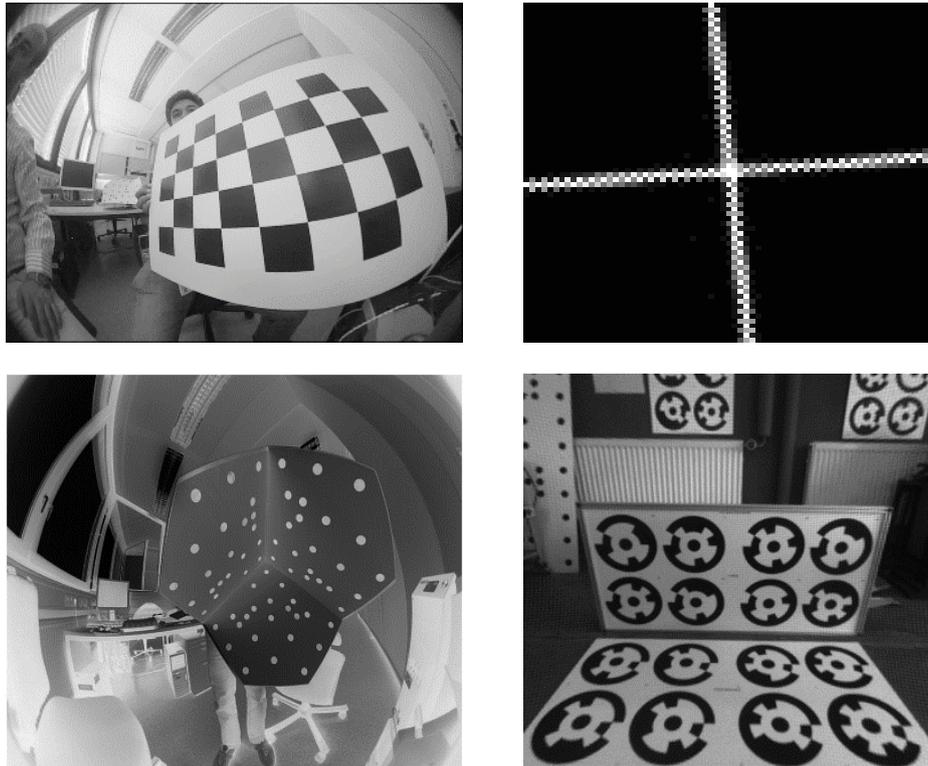
 Fig. 2.29. Test-stand for calibrating central-projection/perspective camera
a: View of a plumb-line calibration test-field , Illustrated after (Fryer & Brown, 1986)
b, c: Illustration of distortion effects on the lines, Illustrated after (Fryer & Brown, 1986)

This approach has also been continuously developed and performed for different digital cameras and benefits a certain ease-of-use in special cases with existing linear and straight features, perpendicular or parallel lines visible in the image (Prescott & McLean, 1997) (Ying, et al., 2006) (Fuqiang, et al., 2013). The limitation of this approach is that mostly the radial distortion could be significantly measured. Other relatively small non-radial distortion effects that do not affect the linearity of the world-to-camera projection are often not considered. Therefore this may not be the most suitable solution when aiming to model all the possible optical effects and deformations.

Recently there have been new calibration methods introduced, which are benefiting from certain geometrical information in object/image space for performing camera calibration such as similarities in objects while considering a unified camera motion (Censi & Scaramuzza, 2013). These methods benefit from the possibility of being performed in an online scenario. However, one drawback is that the existence of those certain geometrical features and their non-deformation assumption in the object space could not be guaranteed. Having no control over the qualities of those geometrical information would directly affect the quality of the camera calibration. Furthermore, these methods require relatively high processing power to perform in real-time and achieve the required accuracies using multiple images. Therefore they are not very suitable for the current mainstream hardware used at the time in most of the driver-assistant systems.

2.6. Image Marker Measurements and Estimation

In every camera calibration solutions today, we have to use different image measurement techniques to find certain geometrical features in image. This information is one of the inputs of each and every calibration solution. Having accurate image measurement is important because it can directly influence the accuracy of camera calibration. There are varieties of algorithms and techniques which could be used in order to measure specific objects and detect patterns in image from which a few are discussed.



a	b
c	d

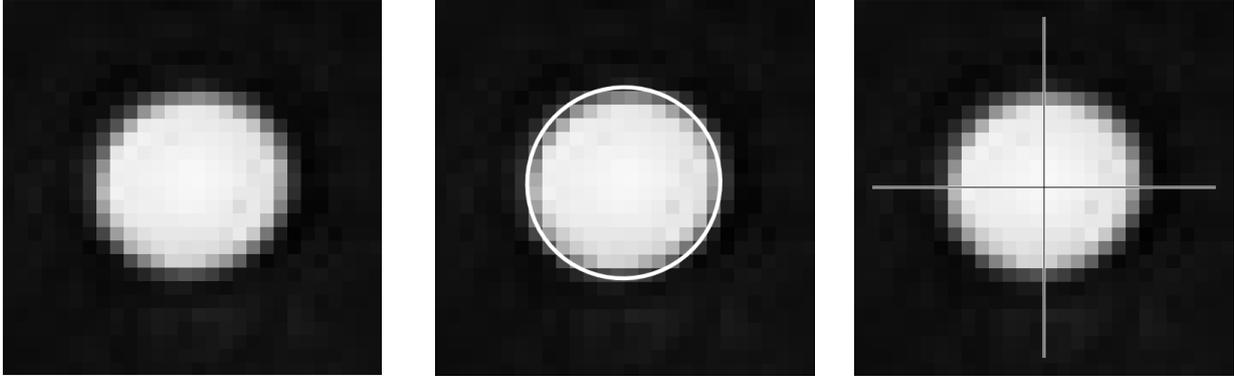
Fig. 2.30. Illustration of marker detection techniques
 a: Checkerboard corner detection for a fisheye camera, b: Slanted crosshair detection e.g. for camera sharpness testing
 c: Circular markers in bundle-block calibration for a fisheye camera. (Courtesy of Robert Bosch GmbH)
 d: Single-coded target points used for bundle-block calibration, PhotoModeler, Eos Systems Inc.

Fig. 2.30 illustrates few different image measurement techniques used in camera calibration. We can see methods such as finding the joint-corners of the checkerboard e.g. for some bundle-block calibration solutions, or determining the centre of a crosshair in case of collimator calibration or the detection of the centre of a circle or ellipse which is the usual form of a target point. The latter is realized mostly in close-range camera calibration using circular markers in the scene (e.g. target points with given 3D coordinates). We should also mention the use of coded targets used for accurate and automatic point matching between images in bundle-block calibration (Niederöst & Maas, 1997) (Schneider, 1996) (Ahn & Kotowski, 1997) (Luhmann, et al., 2011), a technique which is often used in many commercial photogrammetry software.

In most of the automatic calibration processes there are different marker-detection algorithms used to detect circular markers in image with great accuracies. These point-detection algorithms are discussed in the next chapter.

2.6.1. Contour-Track Marker Detection

In Contour-Track marker detection we try to find the contours that represent the edge of the markers in the image. Using these contours we can estimate the centre of the markers in the image. (Luhmann, et al., 2011)



a	b	c
---	---	---

Fig. 2.31. Contour-track marker detection

a: Example of a visible marker in image

b: Approximation of the theoretical contour around the (ideal) circular marker

c: Estimated centre from the detected contours (marker edges)

(Courtesy of Robert Bosch GmbH)

We found the accuracy of contour tracking marker detection to be practically around 0.25 to 0.5 [Pixel] based on several edge detection parameters, which can affect the accuracy and quality of the detected contours such as image quality, noise, sharpness, etc.

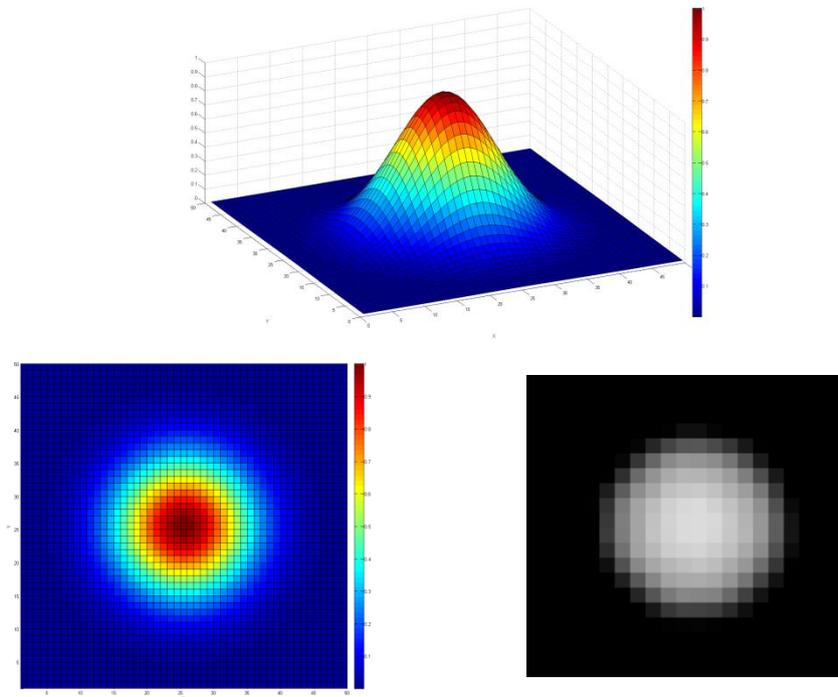
There are different techniques for performing contour detection from which we can name e.g. “profile-driven contour following” or “contour following by gradient analysis” (Luhmann, et al., 2011). Different solutions estimate the centre of target points based on e.g. a detected profile using edge-based analysis or the centre of gravitation of the contour. There are certain parameters, which are important in finding and classifying optimal contours around a target point. From these we can name a few such as the minimum and maximum radius of the targets in image, the in-image calculated values for contour form-factor ($\frac{\text{Circumference}}{\text{Area} \cdot 4 \cdot \pi}$) and the contour concavity factor (Abraham & Hau, 1997) (Krüger & Meinel, 2008) (Payne, et al., 1989) depending on the form of projected targets in image. This is similar to the concept of analysis of curved/concave mirrors (Fritzpatrick, 2007).

The accuracy of contour-track marker detection is usually enough for applications that do not demand very high accuracies below 0.25 [Pixel]. If higher accuracies is needed, the results of contour tracking marker detection i.e. image coordinates of the marker centre point, could be used as input information for other marker detection solutions which require initial coordinates and can deliver accuracies higher than 0.25 [Pixel]. These solutions such as sub-pixel pattern cross-correlation and ellipse-fitting are discussed in the next chapters.

2.6.2. Sub-Pixel Pattern Cross-Correlation

Another solution to marker detection problem is by correlating in-image-space defined patterns and templates (Luhmann, et al., 2011) very similar to the projected markers in image. There is also the possibility to dynamically represent the geometric form of the centrally symmetric markers with different functions instead of using pre-defined marker patterns (Nekouei Shahraki, 2012).

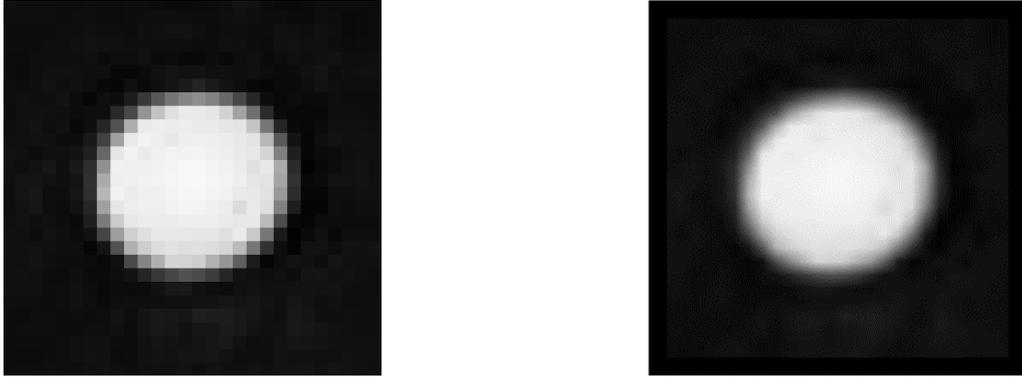
The marker points are usually small white dots on black background. Therefore they could be considered as pulses in the image. In other words one can consider the projected markers in the image as two dimensional centrally symmetrical light pulses. Therefore we can find the position of the markers by cross correlating⁸ the image with two dimensional Gaussian masks with the size (three sigma) equal to the radius of each marker in image. The point of maximum correlation is considered the centre of the target point. For performance improvement, we can use the initial values of point coordinates using other methods such as contour-tracking maker detection [2.6.1].



a	Fig. 2.32. Marker simulation based on templates and known functions a: Three dimensional view of the Gaussian distribution function (Amplitude= 1) b: Two dimensional view of the Gaussian distribution function (Amplitude= 1) c: Illustration of an actual target point visible in image (taken under controlled exposure time interval to avoid over saturation) (Courtesy of Robert Bosch GmbH)
b	
c	

Furthermore, we can consider the fact that in the reality (in the world/scene) there have been more information and higher details, or the so-called high frequency data, that we were not able to record because of a limited capturing frequency i.e. the existing image resolution, which is interpreted as having aliasing. Therefore the sub-pixel sampling could be accompanied by anti-aliasing filters such as the ones used in computer graphics which help us extract more image information and thus get better results (Luhmann, et al., 2011) (Nekouei Shahraki, 2012).

⁸ Generally, the phrase “cross-correlation” is used when performing the correlation of two different functions.



a **B** Fig. 2.33. Sub-pixel sampling and Sub-pixel interpolation
 a: Example of a visible marker in image
 b: Sub-pixel sampling using Gaussian anti-aliasing filter, Sub-Pixel depth = 0.1 [Pixel]
 (Courtesy of Robert Bosch GmbH)

One can usually witness accuracy improvements by using sub-pixel sampling up to 0.1 [Pixel] (Luhmann, et al., 2011) (Nekouei Shahraki, 2012). Higher sampling levels usually in many cases do not show any considerable improvement because at these levels there is not much image information to retrieve.

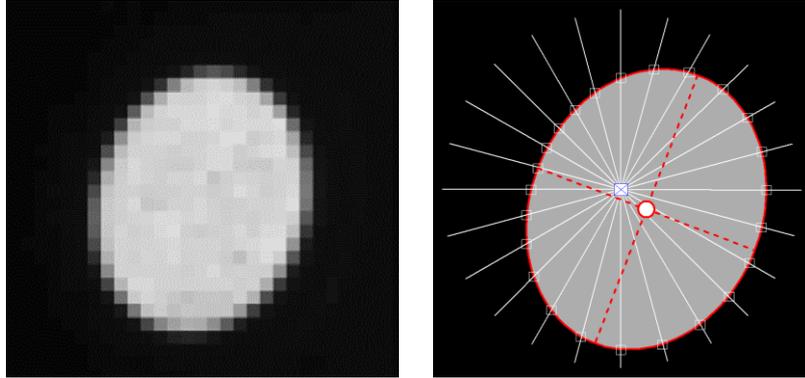


a **b** Fig. 2.34. illustration of the results of different marker detection solutions
 a: Estimated centre from contour tracking marker detection
 b: Estimated centre from sub-pixel pattern cross correlation, Depth= 0.1 [Pixel]
 (Courtesy of Robert Bosch GmbH)

As we can see in Fig. 2.34 the sub-pixel pattern cross correlation has taken the estimated centre from contour tracking marker detection and has improved the coordinates of the marker centre. This solution usually brings more stability to the marker detection process too (Nekouei Shahraki, 2012) because it is less sensitive to noise or a badly visible contour (marker edge) especially for smaller marker sizes. The drawback of this solution is that mostly a high processing power and computational effort is needed (Luhmann, et al., 2011) in order to perform the sub-pixel cross correlation all over the image in nearly run-time. Therefore usually parallel processing approaches such as multi-threaded processing or GPGPU techniques are necessary to boost the performance of this solution.

2.6.3. Ellipse Fitting Marker Detection

The ellipse fitting is a marker detection solution based on fitting an ellipse to the edge of the markers that employs the least squares estimation system. By using this technique also known as least squares matching, one can match the edge of the ellipse to the edge of the markers, and finally consider the centre of the fitted ellipse as the marker centre. The precision of this solution can theoretically reach up to 0.02 [Pixel], which corresponds to the standard deviation of the centre point of ellipse after performing the least-squares estimation (Luhmann, et al., 2011).



a	b
---	---

 Fig. 2.35. Visualisation of ellipse-fitting marker detection (Illustrated after Luhmann, et al., 2011, Fig. 5.62)
a: Visible marker in the image, b: Star operator: fitting an ellipse to the points of intersection between the marker edge and the rays from the first approximation of centre

One of the main advantages of this model – besides having high precision - is that it can still find the centre of the markers accurately in case the circular markers are projected in to the camera in an ellipse form. This happens for example when markers are geometrically placed on a surface not perpendicular to the optical axis or when the marker geometries are affected by the general angular lens behaviour and the reduction of entrance pupil based on the sight angle. This effect is illustrated in Fig. 2.2.

Ellipse-fitting is a suitable target measurements solution for many applications, but also has slight drawbacks. For example when using images with high radial and non-radial distortion usually at the edge of image, the effect of the distortion on the inner-edge of the points is lower than the effect of distortion on the outer-edge. Therefore the shape of the image points deviate systematically from an ellipse form, and we see a decrease in the accuracy of ellipse fitting in these image areas. The recent studies have shown that for a usual perspective camera, the accuracy of the ellipse-fitting is comparable to the template fitting or the sub-pixel cross-correlation if optical distortion is not considered in the marker detection solution (Reznicek, et al., 2016).

We have reviewed the three mostly-used marker detection techniques in this chapter. Similar to what we have discussed, in our tests we use the contour-tracking solution as the initial marker measurement. Then to gain more accuracy, we use the initial coordinates for the ellipse-fitting together with sub-pixel pattern cross-correlation as the two main image measurement techniques.

3. Extended Omnidirectional Camera Model

As discussed in chapter 2.3, fisheye cameras are used for many photogrammetric, computer-vision and driver assistant applications, and therefore one part of this study is focused on the calibration of fisheye cameras. In this chapter, we further development and extend the projection system of Christopher Mei model [2.4.2.5] (Mei & Rives, 2007). We chose this model as the basis of our development work because it generally provides a good modelling accuracy with the capability of fitting to different optical designs as discussed in [2.4.2.5]. The extensions include parametrisation of the projection in traditional photogrammetric format, function linearization, stepwise simplifications and runtime optimizations, which allows modelling the behaviour of fisheye lenses with any opening angles that are used in variety of real-time applications.

We use the C. Mei model to derive the unified projection functions and their linearized derivatives which could later be used - similar to the collinearity equation - in linearized least squares equations and analytical camera self-calibration (Amini, 2006) (Luhmann, et al., 2011). This calibration model has certain advantages, which provides the motivation to further study and extend it. The advantage of developing our model based on C. Mei model is that instead of assuming purely mathematical projection or parabolic mirrors (like some omnidirectional models [2.4.2]), this model assumes a physical projection. Furthermore, there is absolutely no singularity at 180 degrees and no limitation for the opening angle in contrast to some calibration models [2.4.2]. We start by physically modelling the projected rays and model the behaviour of the transparent fish-eye multi-lens system using coupled ray-projections. Therefore, we can practically extract physical camera features from this model.

As extensions and expansions to this fisheye model, considering the projection function, direct object-to-image transformations and inverse image-to-object transformations will be derived using multiple real-time optimizations. Thus, we can both perform the correction to have undistorted image for many applications, and also to use the fish-eye coordinates directly in our calculations. Furthermore, to perform LSE in classical photogrammetric style, the linearized form of the projection formula is derived. The formulas are very similar to the collinearity equation with an extra correction factor. This similarity would later help us a lot in applying the derived equations and interpreting the parameters. Using the extended omnidirectional camera model motivated us to continue with a major accuracy improvement as we later add a more flexible distortion model to the calibration model to increase the accuracy of distortion modelling [4].

3.1. Main Omnidirectional Projection Model

In order to build-up the projection model suitable for omnidirectional cameras, considering the projection behaviour, we first describe the external and internal optical path in geometrical form. After that we combine the equations to derive a generalized projection equation.

3.1.1. Internal and External Lens Modelling (Projections)

As discussed in the previous chapter, similar to the C. Mei model, we use two separate projection steps and then combine them to get the final projection model. As the first projection, we project the object point P (X_p, Y_p, Z_p) on the surface of a sphere with an arbitrary radius R_s and position of centre point CS (X_{CS}, Y_{CS}, Z_{CS}) as projection centre and axes

direction angles (ω, φ, χ):

$$\begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \frac{R_s}{|\vec{r}_{pCS}|} [R_{\omega\varphi\chi}] \begin{bmatrix} X_p - X_{CS} \\ Y_p - Y_{CS} \\ Z_p - Z_{CS} \end{bmatrix}$$

Eq. 3.1

$$|\vec{r}_{pCS}| = \sqrt{(X_p - X_{CS})^2 + (Y_p - Y_{CS})^2 + (Z_p - Z_{CS})^2}$$

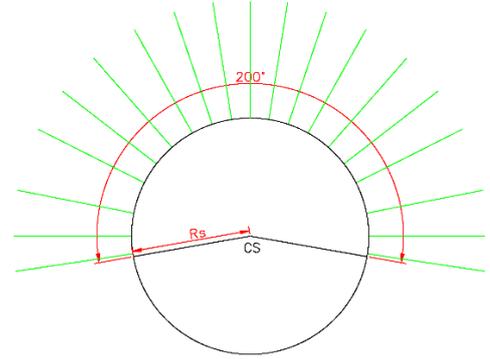


Fig. 3.1. Projection of object points on the surface of sphere with (arbitrary) opening angle

Next, we project the sphere surface points into a virtual standard pinhole camera with projection centre CV (X_{CV}, Y_{CV}, Z_{CV}) and direction angles ($\underline{\Omega}, \underline{\Phi}, \underline{K}$) which results in point p (x_p, y_p, z_p) with the projection scale factor S:

$$\begin{bmatrix} x_p - x_o \\ y_p - y_o \\ z_p - z_o \end{bmatrix} = S [R_{\underline{\Omega}\underline{\Phi}\underline{K}}] \begin{bmatrix} X_s - X_{CV} \\ Y_s - Y_{CV} \\ Z_s - Z_{CV} \end{bmatrix} \quad @ Z_v = -f$$

Eq. 3.2

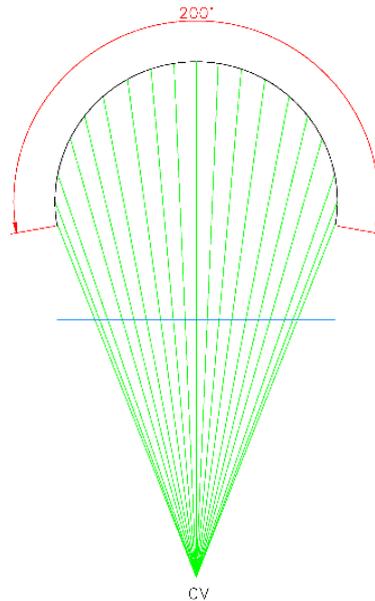


Fig. 3.2. Projection of sphere surface points in a standard pinhole camera

3.1.2. Projection Coupling

Now we combine the two separate projections from Eq. 3.1 and Eq. 3.2 to derive the final omnidirectional model. For simplification purposes, we can set the axes of sphere and the camera coordinates system to be parallel. This implies that for example the angles $(\omega, \varphi, \kappa)$ be equal to the camera direction angles and the $(\underline{\Omega}, \underline{\Phi}, \underline{K})$ should be equal to zero ($R_{\Omega\Phi K} = I$) because no more rotations are necessary.

$$\begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} = S \begin{bmatrix} X_s - X_{CV} \\ Y_s - Y_{CV} \\ Z_s - Z_{CV} \end{bmatrix} \quad \text{Eq. 3.3}$$

Eq. 3.3 gives the second projection for a specific internal alignment. Again for simplification purposes, we can set the (X_{CV}, Y_{CV}, Z_{CV}) equal to $(0, 0, f + H_v)$. This means the second projection centre lies on the optical axis with the coordinates system origin is CS. This gives the equivalent optical distance of the second projection:

$$\begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} = S \begin{bmatrix} X_s \\ Y_s \\ Z_s - (f + H_v) \end{bmatrix} \quad \text{Eq. 3.4}$$

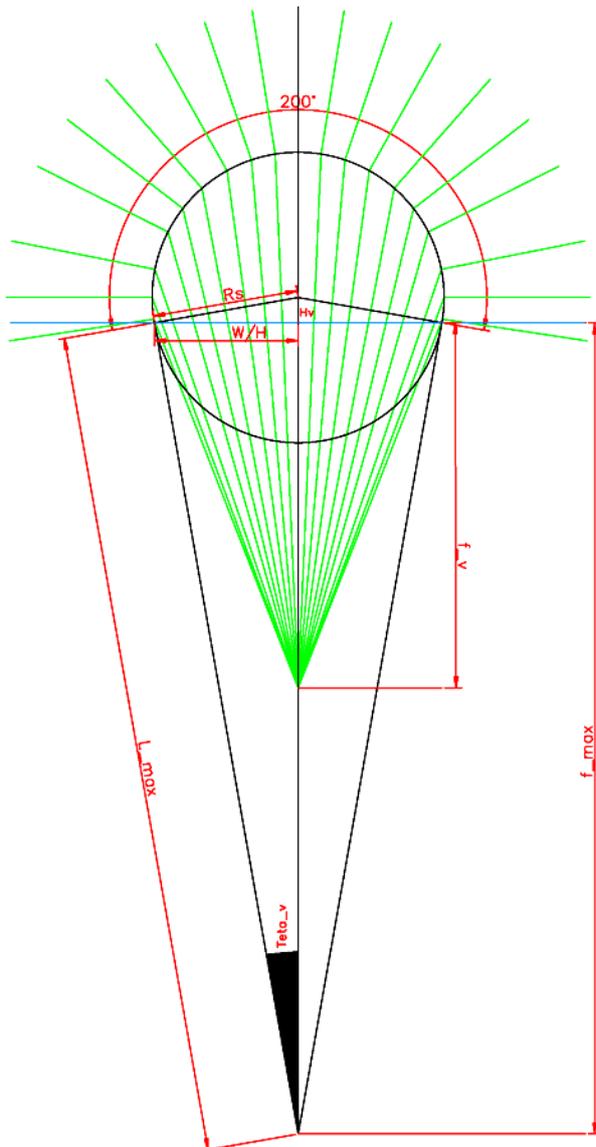


Fig. 3.3. Coupling of the two projections together to create a unified omnidirectional projection model (Courtesy of Robert Bosch GmbH)

By coupling the two projections, we would have:

$$\begin{cases} \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \frac{1}{|\vec{r}_{pCS}|} [R_{\omega\phi\kappa}] \begin{bmatrix} X_P - X_{CS} \\ Y_P - Y_{CS} \\ Z_P - Z_{CS} \end{bmatrix} \\ \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} = S \begin{bmatrix} X_s \\ Y_s \\ Z_s - (f + H_v) \end{bmatrix} \end{cases} @ |\vec{r}_{pCS}| = \sqrt{(X_P - X_{CS})^2 + (Y_P - Y_{CS})^2 + (Z_P - Z_{CS})^2} \quad \text{Eq. 3.5}$$

This will give us the unified form of the coupled projections describing the general fisheye projection:

$$\Rightarrow \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} = S \left(\frac{1}{|\vec{r}_{pCS}|} [R_{\omega\phi\kappa}] \begin{bmatrix} X_P - X_{CS} \\ Y_P - Y_{CS} \\ Z_P - Z_{CS} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ (f + H_v) \end{bmatrix} \right) \quad \text{Eq. 3.6}$$

The above projection will further be written in unified formalism and linearized in the following chapters.

3.2. Direct Transformation: World to Camera (W2C)

In this step, we derive the projection equation from object space to image space based on the coupled projection in the previous chapter:

$$\begin{aligned} \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} &= S \left(\frac{1}{|\vec{r}_{pCS}|} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} X_P - X_{CS} \\ Y_P - Y_{CS} \\ Z_P - Z_{CS} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ (f + H_v) \end{bmatrix} \right) \\ \Rightarrow \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} &= \frac{S}{|\vec{r}_{pCS}|} \left(\begin{bmatrix} m_{11}(X_P - X_{CS}) + m_{12}(Y_P - Y_{CS}) + m_{13}(Z_P - Z_{CS}) \\ m_{21}(X_P - X_{CS}) + m_{22}(Y_P - Y_{CS}) + m_{23}(Z_P - Z_{CS}) \\ m_{31}(X_P - X_{CS}) + m_{32}(Y_P - Y_{CS}) + m_{33}(Z_P - Z_{CS}) \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ |\vec{r}_{pCS}|(f + H_v) \end{bmatrix} \right), \quad \text{Eq. 3.7} \\ \Rightarrow \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} &= \frac{S}{|\vec{r}_{pCS}|} \begin{bmatrix} m_{11}(X_P - X_{CS}) + m_{12}(Y_P - Y_{CS}) + m_{13}(Z_P - Z_{CS}) \\ m_{21}(X_P - X_{CS}) + m_{22}(Y_P - Y_{CS}) + m_{23}(Z_P - Z_{CS}) \\ m_{31}(X_P - X_{CS}) + m_{32}(Y_P - Y_{CS}) + m_{33}(Z_P - Z_{CS}) - |\vec{r}_{pCS}|(f + H_v) \end{bmatrix} \end{aligned}$$

Eq. 3.7 gives the formalism of the coupled fisheye projection in form of a single world-to-camera projection. Now in order to be able to perform analytical linear algebra calculations, we need to first linearize the equation system. This is discussed in the next chapter.

3.3. Linearization (Tailor series derivatives)

In this chapter we perform the linearization of the projection equation derived in the previous chapter. This helps us to use the formalism in our least squares equation system. In order to setup the equations, we perform the linearization for the direct world-to-camera transformation (W2C). The projection equation then reads:

$$\text{Vector of Unknowns (Exterior Orientation)} = \begin{bmatrix} \omega \\ \varphi \\ \chi \\ X_{CS} \\ Y_{CS} \\ Z_{CS} \end{bmatrix} \quad \text{(Rotation Matrix)} \quad M_{\chi\varphi\omega} = M_{\chi}M_{\varphi}M_{\omega} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

Eq. 3.8

$$\begin{cases} x_p - x_o = -f \frac{m_{11}(X_P - X_{CS}) + m_{12}(Y_P - Y_{CS}) + m_{13}(Z_P - Z_{CS})}{m_{31}(X_P - X_{CS}) + m_{32}(Y_P - Y_{CS}) + m_{33}(Z_P - Z_{CS}) - |\vec{r}_{PCS}|(f + H_v)} \\ y_p - y_o = -f \frac{m_{21}(X_P - X_{CS}) + m_{22}(Y_P - Y_{CS}) + m_{23}(Z_P - Z_{CS})}{m_{31}(X_P - X_{CS}) + m_{32}(Y_P - Y_{CS}) + m_{33}(Z_P - Z_{CS}) - |\vec{r}_{PCS}|(f + H_v)} \end{cases}$$

In Eq. 3.8, the formalism of the fisheye projection function similar to the collinearity equation is given. The x_o, y_o coordinates are the interior orientation parameters or the coordinates of the principle point (PPS). We define the following notations for simplifications and compatibility with classical photogrammetric notations:

$$\begin{cases} F = x_o - x_p - f \frac{r}{q} = 0 \\ G = y_o - y_p - f \frac{s}{q} = 0 \end{cases}$$

$$\begin{cases} q = m_{31}(\Delta X) + m_{32}(\Delta Y) + m_{33}(\Delta Z) - |\vec{r}_{PCS}|(f + H_v) \\ r = m_{11}(\Delta X) + m_{12}(\Delta Y) + m_{13}(\Delta Z) \\ s = m_{21}(\Delta X) + m_{22}(\Delta Y) + m_{23}(\Delta Z) \end{cases} \quad \begin{cases} \Delta X = X_P - X_{CS} \\ \Delta Y = Y_P - Y_{CS} \\ \Delta Z = Z_P - Z_{CS} \end{cases}$$

Eq. 3.9

$$\begin{cases} F = F_0 + \left(\frac{\partial F}{\partial x_p}\right) dx_p + \left(\frac{\partial F}{\partial \omega}\right) d\omega + \left(\frac{\partial F}{\partial \varphi}\right) d\varphi + \left(\frac{\partial F}{\partial \chi}\right) d\chi + \left(\frac{\partial F}{\partial X_{CS}}\right) dX_{CS} + \left(\frac{\partial F}{\partial Y_{CS}}\right) dY_{CS} + \left(\frac{\partial F}{\partial Z_{CS}}\right) dZ_{CS} \\ \quad + \left(\frac{\partial F}{\partial X_p}\right) dX_p + \left(\frac{\partial F}{\partial Y_p}\right) dY_p + \left(\frac{\partial F}{\partial Z_p}\right) dZ_p + \left(\frac{\partial F}{\partial f}\right) df + \left(\frac{\partial F}{\partial H_v}\right) dH_v = 0 \\ G = G_0 + \left(\frac{\partial G}{\partial y_p}\right) dy_p + \left(\frac{\partial G}{\partial \omega}\right) d\omega + \left(\frac{\partial G}{\partial \varphi}\right) d\varphi + \left(\frac{\partial G}{\partial \chi}\right) d\chi + \left(\frac{\partial G}{\partial X_{CS}}\right) dX_{CS} + \left(\frac{\partial G}{\partial Y_{CS}}\right) dY_{CS} + \left(\frac{\partial G}{\partial Z_{CS}}\right) dZ_{CS} \\ \quad + \left(\frac{\partial G}{\partial X_p}\right) dX_p + \left(\frac{\partial G}{\partial Y_p}\right) dY_p + \left(\frac{\partial G}{\partial Z_p}\right) dZ_p + \left(\frac{\partial G}{\partial f}\right) df + \left(\frac{\partial G}{\partial H_v}\right) dH_v = 0 \end{cases}$$

The linearization of the equations of the extended omnidirectional model is given in Eq. 3.9, similar to the linearization process of collinearity equation.

The following notations are used for simplifications and compatibility with classical photogrammetric notations:

$$\begin{aligned}
 v_{x_p} &= dx_p \quad \text{and} \quad \frac{\partial F}{\partial x_p} = -1 \\
 \Rightarrow v_{x_p} &= \left(\frac{\partial F}{\partial \omega}\right) d\omega + \left(\frac{\partial F}{\partial \varphi}\right) d\varphi + \left(\frac{\partial F}{\partial \kappa}\right) d\kappa + \dots + F_0 \\
 v_{y_p} &= dy_p \quad \text{and} \quad \frac{\partial G}{\partial y_p} = -1 \\
 \Rightarrow v_{y_p} &= \left(\frac{\partial G}{\partial \omega}\right) d\omega + \left(\frac{\partial G}{\partial \varphi}\right) d\varphi + \left(\frac{\partial G}{\partial \kappa}\right) d\kappa + \dots + G_0
 \end{aligned} \tag{Eq. 3.10}$$

With respect to the x coordinates (or F equation), using initial values, we can calculate the derivatives of the projection function for linearization:

$$\begin{aligned}
 v_{x_p} &= b_{11}d\omega + b_{12}d\varphi + b_{13}d\kappa - b_{14}dX_{CS} - b_{15}dY_{CS} - b_{16}dZ_{CS} + b_{14}dX_P + b_{15}dY_P \\
 &\quad + b_{16}dZ_P + J \\
 @ \quad b_{1i} &= \frac{\partial F}{\partial u} = \frac{f}{q^2} \left(r \frac{\partial q}{\partial u} - q \frac{\partial r}{\partial u} \right) \quad @ \quad i = \text{index of parameter } u \quad @ \quad J = F_0 = \text{cte}
 \end{aligned} \tag{Eq. 3.11}$$

Then we have:

$$b_{11} = \frac{\partial F}{\partial \omega} = \frac{f}{q^2} [r(-m_{33}\Delta Y + m_{32}\Delta Z) - q(-m_{13}\Delta Y + m_{12}\Delta Z)] \tag{Eq. 3.12}$$

$$\begin{aligned}
 b_{12} = \frac{\partial F}{\partial \varphi} &= \frac{f}{q^2} [r(\cos(\varphi)\Delta X + \sin(\omega)\sin(\varphi)\Delta Y - \cos(\omega)\sin(\varphi)\Delta Z) \\
 &\quad - q(-\sin(\varphi)\cos(\kappa)\Delta X + \sin(\omega)\cos(\varphi)\cos(\kappa)\Delta Y - \cos(\omega)\cos(\varphi)\cos(\kappa)\Delta Z)]
 \end{aligned} \tag{Eq. 3.13}$$

$$b_{13} = \frac{\partial F}{\partial \kappa} = -\frac{f}{q} (m_{21}\Delta X + m_{22}\Delta Y + m_{23}\Delta Z) \tag{Eq. 3.14}$$

The derivatives b_{11}, b_{12}, b_{13} are now calculated.

Now we calculate b_{14}, b_{15}, b_{16} which are the derivation of the $|\vec{r}_{pCS}|$ with respect to the camera position. Recalling that:

$$\begin{aligned}
 |\vec{r}_{pCS}| &= \sqrt{(X_P - X_{CS})^2 + (Y_P - Y_{CS})^2 + (Z_P - Z_{CS})^2} \\
 \Rightarrow \frac{\partial}{\partial X_P} |\vec{r}_{pCS}| &= -\frac{\partial}{\partial X_{CS}} |\vec{r}_{pCS}| = \frac{2(X_P - X_{CS})}{2|\vec{r}_{pCS}|} \\
 \Rightarrow \frac{\partial}{\partial Y_P} |\vec{r}_{pCS}| &= -\frac{\partial}{\partial Y_{CS}} |\vec{r}_{pCS}| = \frac{2(Y_P - Y_{CS})}{2|\vec{r}_{pCS}|} \\
 \Rightarrow \frac{\partial}{\partial Z_P} |\vec{r}_{pCS}| &= -\frac{\partial}{\partial Z_{CS}} |\vec{r}_{pCS}| = \frac{2(Z_P - Z_{CS})}{2|\vec{r}_{pCS}|}
 \end{aligned} \tag{Eq. 3.15}$$

Therefore:

$$\begin{aligned}
 b_{14} &= \frac{\partial F}{\partial X_P} = \frac{f}{q^2} \left(r \frac{\partial q}{\partial X_P} - q \frac{\partial r}{\partial X_P} \right) = \frac{f}{q^2} \left(r(m_{31} - \frac{2(X_P - X_{CS})(f + H_v)}{2|\vec{r}_{pCS}|}) - qm_{11} \right) \\
 b_{15} &= \frac{\partial F}{\partial Y_P} = \frac{f}{q^2} \left(r \frac{\partial q}{\partial Y_P} - q \frac{\partial r}{\partial Y_P} \right) = \frac{f}{q^2} \left(r(m_{32} - \frac{2(Y_P - Y_{CS})(f + H_v)}{2|\vec{r}_{pCS}|}) - qm_{12} \right) \\
 b_{16} &= \frac{\partial F}{\partial Z_P} = \frac{f}{q^2} \left(r \frac{\partial q}{\partial Z_P} - q \frac{\partial r}{\partial Z_P} \right) = \frac{f}{q^2} \left(r(m_{33} - \frac{2(Z_P - Z_{CS})(f + H_v)}{2|\vec{r}_{pCS}|}) - qm_{13} \right)
 \end{aligned} \tag{Eq. 3.16}$$

Now we also add the extra self-calibration parameters:

$$\begin{aligned}
 b_{17} &= \frac{\partial F}{\partial f} = -\frac{r}{q} + \frac{f}{q^2} \left(r \frac{\partial q}{\partial f} - q \frac{\partial r}{\partial f} \right) = -\frac{r}{q} - \frac{f}{q^2} (r |\vec{r}_{pCS}|) \\
 b_{18} &= \frac{\partial F}{\partial H_v} = \frac{f}{q^2} \left(r \frac{\partial q}{\partial H_v} - q \frac{\partial r}{\partial H_v} \right) = -\frac{f}{q^2} (r |\vec{r}_{pCS}|)
 \end{aligned} \tag{Eq. 3.17}$$

And also similar to the previously derived equations for the x coordinates, we derive the equation with respect to the y coordinates (G equation). Using the initial values, it reads:

$$v_{y_p} = G - G_0 = b_{21}d\omega + b_{22}d\varphi + b_{23}d\chi - b_{24}dX_{CS} - b_{25}dY_{CS} - b_{26}dZ_{CS} + b_{24}dX_P + b_{25}dY_P + b_{26}dZ_P + K$$

$$@ \quad b_{2i} = \frac{\partial G}{\partial u} = \frac{f}{q^2} \left(s \frac{\partial q}{\partial u} - q \frac{\partial s}{\partial u} \right) \quad @ i = \text{index of parameter "u"} \quad @ K = G_0 = \text{cte}$$
Eq. 3.18

Then we have:

$$b_{21} = \frac{\partial G}{\partial \omega} = \frac{f}{q^2} [s(-m_{33}\Delta Y + m_{32}\Delta Z) - q(-m_{23}\Delta Y + m_{22}\Delta Z)]$$

$$b_{22} = \frac{\partial G}{\partial \varphi} = \frac{f}{q^2} [s(\cos(\varphi)\Delta X + \sin(\omega)\sin(\varphi)\Delta Y - \cos(\omega)\sin(\varphi)\Delta Z) - q(\sin(\varphi)\sin(\chi)\Delta X - \sin(\omega)\cos(\varphi)\sin(\chi)\Delta Y + \cos(\omega)\cos(\varphi)\sin(\chi)\Delta Z)]$$

$$b_{23} = \frac{\partial G}{\partial \chi} = \frac{f}{q} (m_{11}\Delta X + m_{12}\Delta Y + m_{13}\Delta Z)$$
Eq. 3.19

The Coefficients b_{21}, b_{22}, b_{23} are calculated. Now calculating b_{24}, b_{25}, b_{26} :

$$b_{24} = \frac{\partial G}{\partial X_P} = \frac{f}{q^2} \left(s \frac{\partial q}{\partial X_P} - q \frac{\partial s}{\partial X_P} \right) = \frac{f}{q^2} \left(s(m_{31} - \frac{2(X_P - X_{CS})(f + H_v)}{2|\vec{r}_{PCS}|}) - qm_{21} \right)$$

$$b_{25} = \frac{\partial G}{\partial Y_P} = \frac{f}{q^2} \left(s \frac{\partial q}{\partial Y_P} - q \frac{\partial s}{\partial Y_P} \right) = \frac{f}{q^2} \left(s(m_{32} - \frac{2(Y_P - Y_{CS})(f + H_v)}{2|\vec{r}_{PCS}|}) - qm_{22} \right)$$

$$b_{26} = \frac{\partial G}{\partial Z_P} = \frac{f}{q^2} \left(s \frac{\partial q}{\partial Z_P} - q \frac{\partial s}{\partial Z_P} \right) = \frac{f}{q^2} \left(s(m_{33} - \frac{2(Z_P - Z_{CS})(f + H_v)}{2|\vec{r}_{PCS}|}) - qm_{23} \right)$$
Eq. 3.20

Now we add the extra self-calibration parameters:

$$b_{27} = \frac{\partial G}{\partial f} = -\frac{s}{q} + \frac{f}{q^2} \left(s \frac{\partial q}{\partial f} - q \frac{\partial s}{\partial f} \right) = -\frac{s}{q} - \frac{f}{q^2} (s |\vec{r}_{PCS}|)$$

$$b_{28} = \frac{\partial G}{\partial H_v} = \frac{f}{q^2} \left(s \frac{\partial q}{\partial H_v} - q \frac{\partial s}{\partial H_v} \right) = -\frac{f}{q^2} (s |\vec{r}_{PCS}|)$$
Eq. 3.21

3.4. Setting-Up Equations System

After performing the linearization of the projection equations in the previous chapter, and having all the b_{ij} coefficients, we can build up the LSE equations system or generally perform linear algebra calculations.

$$\underbrace{\begin{bmatrix} v_{x_p} \\ v_{y_p} \\ \vdots \end{bmatrix}}_{dL} = \underbrace{\begin{bmatrix} b_{11} & b_{12} & b_{13} & -b_{14} & -b_{15} & -b_{16} & b_{17} & b_{18} \\ b_{21} & b_{22} & b_{23} & -b_{24} & -b_{25} & -b_{26} & b_{27} & b_{28} \\ \vdots & \vdots \end{bmatrix}}_A \underbrace{\begin{bmatrix} d\omega \\ d\varphi \\ dx \\ dX_{CS} \\ dY_{CS} \\ dZ_{CS} \\ df \\ dH_v \\ dx \end{bmatrix}}_{dX} \quad \text{Eq. 3.22}$$

$$dL = A \cdot dX$$

In Eq. 3.22, we see the set-up of the linearized least squares equation system using the projection derivations, and in grey colour are the parameters for optional estimation of parameter H_v . All the derivatives are taken from the corresponding equations or estimated using numerical derivations. As shown in Eq. 3.22 the built up least squares equation system is quite similar to a perspective projection camera and is solved using the equations given in appendix 10.1.

3.4.1. Geometrical parameter correlation

Generally, when performing camera calibration we need to be aware of the possible correlation between the intrinsic and extrinsic parameters. When performing linearized least-squares from non-linear functions, good initial values are required. Also having badly distributed points in space makes the realization and modelling of the first projection not possible and also creates parameter instability. Having too many distortion parameters will introduce over-fitting which results in high correlation between calibration parameters and should be avoided.

As previously discussed, by adding the calibration parameters to our equation system we can perform self-calibration. However, as also mentioned briefly in (Mei & Rives, 2007) we may have high correlation between the projection parameters f and H_v . It can be seen in Eq. 3.17 and Eq. 3.21 that by considering both of the parameters f and H_v as unknown in the equation system, our equation system may experience some correlation instability between these parameters.

$$\begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} = S \left(\frac{1}{|\vec{r}_{pCS}|} [R_{\omega\varphi\lambda}] \begin{bmatrix} X_p - X_{CS} \\ Y_p - Y_{CS} \\ Z_p - Z_{CS} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ H_{v2} \end{bmatrix} \right) \quad \text{Eq. 3.23}$$

In order to solve the correlation problem and to increase the geometrical and numerical stability we can consider the parameter H_v as pre-defined fixed parameter and only estimate the camera focal length f during calibration. Alternatively, for simplification we can reparametrize the Eq. 3.6 to have $H_{v2} = f + H_v$ and again considering H_{v2} as fixed parameter, described in Eq. 3.23.

3.4.2. Distortion model for extended fisheye calibration

As mentioned in [3.4.1], by selecting suitable values for the Parameter H_v in Eq. 3.6 or H_{v2} in Eq. 3.23 a big part of the fish-eye projection can be directly modelled. Therefore, to model the remaining distortion or - in other words - the deviation of optical mapping from the projection function, we can use the familiar distortion models, which are used in pinhole cameras. A suitable distortion model [2.4.3] should be used as additional parameters to model the lens distortion and compensate the small deviations of the camera projection from the ideal initial projection:

$$F = x_o - x_p - f \frac{r}{q} + \text{calibration parameters} \quad \text{Eq. 3.24}$$

As discussed in 2.4.2.5 the distortion model in the C. Mei calibration model is based on the brown distortion model with radial distortion polynomials. In ideal cases when the distortion is not very big and the local lens distortions are small and do not include irregularities, this model can deliver accuracies of up to 0.2 [Pixel].

As a further extension to this calibration model, we have used a distortion model to model the optical lens distortion effects with very high sub-pixel accuracy better than 0.2 [Pixel]. In [4] we introduce the free-function model, which later we use to model different lens distortion.

3.5. Inverse Transformation: Camera to World (C2W)

In this chapter, we model the projection from image space to object space and derive the equations for camera-to-world transformation or the so-called inverse transformation. The C2W transformation is used together with fast image undistortion in order to calculate the direction vectors per pixel. The direction vectors could be used in many machine vision applications. For example by transformation of detected image objects to world objects in object space for 3D reconstruction in systems such as driver assistant systems. The C2W transformation is initially derived from the direct transformation. Therefore, we would proceed as follows:

$$\begin{aligned}
 \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} &= S \left(\frac{1}{|\vec{r}_{pCS}|} [R_{\omega\phi\chi}] \begin{bmatrix} X_p - X_{CS} \\ Y_p - Y_{CS} \\ Z_p - Z_{CS} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ (f + H_v) \end{bmatrix} \right) \\
 \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f \end{bmatrix} &= \frac{S}{|\vec{r}_{pCS}|} [R_{\omega\phi\chi}] \begin{bmatrix} X_p - X_{CS} \\ Y_p - Y_{CS} \\ Z_p - Z_{CS} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ S(f + H_v) \end{bmatrix} \\
 \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f + S(f + H_v) \end{bmatrix} &= \frac{S}{|\vec{r}_{pCS}|} [R_{\omega\phi\chi}] \begin{bmatrix} X_p - X_{CS} \\ Y_p - Y_{CS} \\ Z_p - Z_{CS} \end{bmatrix} & \text{Eq. 3.25} \\
 \frac{|\vec{r}_{pCS}|}{S} [R_{\omega\phi\chi}^T] \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f + S(f + H_v) \end{bmatrix} &= \begin{bmatrix} X_p - X_{CS} \\ Y_p - Y_{CS} \\ Z_p - Z_{CS} \end{bmatrix} \\
 \frac{|\vec{r}_{pCS}|}{S} \begin{bmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \end{bmatrix} \begin{bmatrix} x_p - x_o \\ y_p - y_o \\ -f + S(f + H_v) \end{bmatrix} &= \begin{bmatrix} X_p - X_{CS} \\ Y_p - Y_{CS} \\ Z_p - Z_{CS} \end{bmatrix}
 \end{aligned}$$

In Eq. 3.25, the inverse camera-to-world transformation is setup. The m_{ij} parameters are the elements of the rotation matrix.

$$\boxed{X_p - X_{CS} = (Z_p - Z_{CS}) \frac{m_{11}(x_p - x_o) + m_{21}(y_p - y_o) + m_{31}GF(x_p - x_o, y_p - y_o)}{m_{13}(x_p - x_o) + m_{23}(y_p - y_o) + m_{33}GF(x_p - x_o, y_p - y_o)}}$$

Eq. 3.26

$$\boxed{Y_p - Y_{CS} = (Z_p - Z_{CS}) \frac{m_{12}(x_p - x_o) + m_{22}(y_p - y_o) + m_{32}GF(x_p - x_o, y_p - y_o)}{m_{13}(x_p - x_o) + m_{23}(y_p - y_o) + m_{33}GF(x_p - x_o, y_p - y_o)}}$$

As it could be seen in the formulas in Eq. 3.26, the inverse camera-to-world transformation is similar to the formalism of the inverse collinearity equation., For simplifications in the formulas, a generalized virtual function based on the camera focal-distance is defined (GF). When using this function, the formulas will be comparable –in form- to the inverse of collinearity condition equation (Amini, 2006) and could generally be used in least-squares estimation systems such as bundle-block calibration.

In order to derive the linearized equations for inverse transformation we need to calculate GF analytically. Furthermore, we can derive the GF term for different projection cases, which are later discussed in this chapter. For this purpose, this term has to be analytically derived:

$$GF(x_p - x_o, y_p - y_o) = -f + S(f + H_v) \quad @ \quad S = \frac{L_p}{L_s} \quad \text{or} \quad S = \frac{f}{f + H_v + Z_s} \quad @ \quad Z_s > 0 \quad \text{Eq. 3.27}$$

The general focal distance equivalent (GF) and the internal scaling factor (S) are given in Eq. 3.27.

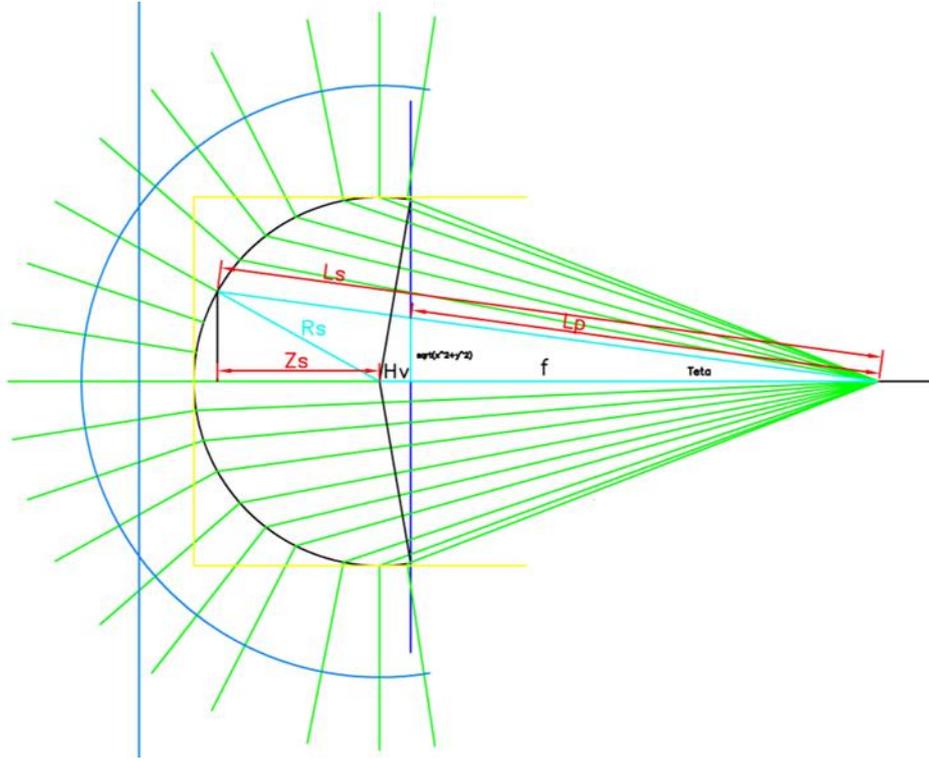


Fig. 3.4. Illustration of the internal parameters of the extended omnidirectional camera projection (Courtesy of Robert Bosch GmbH)

In order to find GF and S, having the projection geometry in mind, we can write:

$$\cos(\vartheta_i) = \frac{f}{\sqrt{f^2 + (x_p - x_o)^2 + (y_p - y_o)^2}} \quad \text{Eq. 3.28}$$

$$R^2 = L_s^2 + (f + H_v)^2 - 2 \cos(\vartheta_i) L_s (f + H_v)$$

$$S = \frac{L_p}{L_s} \Rightarrow \frac{1}{S} = \frac{L_s}{\sqrt{f^2 + (x_p - x_o)^2 + (y_p - y_o)^2}}$$

Using the parameters from Eq. 3.28, we have:

$$\Rightarrow R^2 = \left(\frac{1}{S^2}\right) \left(f^2 + (x_p - x_o)^2 + (y_p - y_o)^2\right) + (f + H_v)^2 - 2 f \left(\frac{1}{S}\right) (f + H_v) \quad \text{Eq. 3.29}$$

$$\Rightarrow S^2((f + H_v)^2 - R^2) + 2S(-f(f + H_v)) + \left(f^2 + (x_p - x_o)^2 + (y_p - y_o)^2\right) = 0$$

Eq. 3.29 describes the calculation of the projection internal scale-factor based on the camera geometrical parameters.

In order to calculate the GF for a generalized projection case, we follow:

$$\begin{aligned}
 & S^2 \underbrace{\left((f + H_v)^2 - R^2 \right)}_a + 2S \underbrace{\left(-f(f + H_v) \right)}_{b'} + \underbrace{\left(f^2 + (x_p - x_o)^2 + (y_p - y_o)^2 \right)}_c = 0 \\
 & S = \frac{-b' - \sqrt{b'^2 - ac}}{a} \quad \text{Having the geometry in mind: } L_p \propto S \\
 & \Rightarrow S = \frac{f(f + H_v) - \sqrt{f^2(f + H_v)^2 - ((f + H_v)^2 - R^2) \left(f^2 + (x_p - x_o)^2 + (y_p - y_o)^2 \right)}}{(f + H_v)^2 - R^2} \quad \text{Eq. 3.30} \\
 & \Rightarrow S = \frac{f(f + H_v) - \sqrt{f^2 R^2 - ((f + H_v)^2 - R^2) \left((x_p - x_o)^2 + (y_p - y_o)^2 \right)}}{(f + H_v)^2 - R^2}
 \end{aligned}$$

And finally the scale factor S is calculated. We know from Eq. 3.27:

$$GF(x_p - x_o, y_p - y_o) = -f + S(f + H_v)$$

Therefore, for the inverse transformation (C2W) we would finally have:

$$\boxed{GF(x_p - x_o, y_p - y_o) = \frac{f R^2 - (f + H_v) \sqrt{f^2 R^2 - ((f + H_v)^2 - R^2) \left((x_p - x_o)^2 + (y_p - y_o)^2 \right)}}{(f + H_v)^2 - R^2}} \quad \text{Eq. 3.31}$$

Calculation of the generalized focal distance (GF) directly from the camera geometrical parameters and image coordinates is given in Eq. 3.31. We should note that f and R and H_v are projection parameters and are all constants. The GF term has now been derived analytically. Using a standard procedure, after calibrating the camera, or generally saying, solving the LSE for direct transformation, we can use the results (the evaluated f and H_v) to calculate GF. Then we would be able to perform online analytical spatial rectification (Amini, 2006) in many applications such as stereo-vision systems, SFM, 3DTV, etc.

3.6. Extracting Physical Parameters and Limits

In iterative least-squares estimation, initial values are required to initialize the linearized equations of non-linear functions. Based on the optical design data we can estimate suitable initial values for the camera model parameters. This also helps us to compare the estimated parameters with the valid limits to be able to evaluate their correctness. In this chapter, we extract some limits for the physical camera parameters from the extended omnidirectional model.

Let's assume the constant camera field-of-view or opening angle $\alpha > \pi$. The distance of the virtual camera plane from CS point would be:

$$H_v = R_s \cos\left(\pi - \frac{\alpha}{2}\right) \quad \text{Eq. 3.32}$$

The distance of the virtual camera from the CS point (H_v) is given in Eq. 3.32. The calculation of the dimension limits of the virtual camera plane is as follows:

$$W = H = R_s \sin\left(\pi - \frac{\alpha}{2}\right) = R_s \sin\left(\frac{\alpha}{2}\right) \quad \text{Eq. 3.33}$$

The limit for selecting the internal focal distance based on the opening angle is derived in Eq. 3.34. This is the limits with which the second projection exists.

$$\begin{aligned} \vartheta_v &= \frac{\alpha}{2} - \frac{\pi}{2}, \quad \frac{W}{f_{\max}} = \tan(\vartheta_v) \\ f_{\max} &= \frac{W}{\tan\left(\frac{\alpha}{2} - \frac{\pi}{2}\right)} = \frac{R_s \sin\left(\frac{\alpha}{2}\right)}{\tan\left(\frac{\alpha}{2} - \frac{\pi}{2}\right)} = \frac{R_s \sin\left(\frac{\alpha}{2}\right)}{-\cot\left(\frac{\alpha}{2}\right)} \\ \Rightarrow f_{\max} &= -R_s \sin\left(\frac{\alpha}{2}\right) \tan\left(\frac{\alpha}{2}\right) \quad f \geq 0 \leftrightarrow \alpha \geq \pi \end{aligned} \quad \text{Eq. 3.34}$$

It is clear that the second projection is applicable only when $\alpha < 2\pi$. Therefore, the selected focal distance should be within this validity domain and satisfy this criterion:

$$0 < f \leq f_{\max} \quad \text{Eq. 3.35}$$

There is an interesting feature when $0 < \alpha < \pi$. In this domain, the focal distance would be negative and the projection plane would lie on the other side of the CS point. Nevertheless, the projection is still fully applicable at this opening angle domain.

3.7. Generalized Projection and Runtime Optimization

In this chapter, we optimize the camera-to-world projection of the extended omnidirectional model. The simplification will help us to simplify the formalism and is beneficial to decrease the calculation time on low-power and mobile platforms. By trying to build up the equations for the angular representation of C2W and W2C projection, we continue to optimize the equation of C2W for runtime calculation.

3.7.1. General Angular W2C Projection

As we had in previous chapters, the parameters H_v , f , α , R_s are projection parameters and are all constants. Therefore we can derive the equations for omnidirectional projection to have a simplified radial or angular-based projection model. As we know from Eq. 3.26, Eq. 3.32 and Eq. 3.34:

$$f = -R_s \sin\left(\frac{\alpha}{2}\right) \tan\left(\frac{\alpha}{2}\right), \quad H_v = R_s \cos\left(\pi - \frac{\alpha}{2}\right) = -R_s \cos\left(\frac{\alpha}{2}\right)$$

$$r_{\text{image}} = S \cdot R_{xy} = S_0 \cdot \frac{f R_{xy}}{f + H_v + Z_s} = \frac{f S_0 R \sin(\varphi)}{f + H_v + R \cos(\varphi)}$$
Eq. 3.36

We define auxiliary parameters a_1 and a_2 :

$$a_1 = f \cdot S_0, \quad a_2 = \frac{f + H_v}{R}$$
Eq. 3.37

Now in case we use unified image coordinates, then we have to set $a_1 = S_0$. Therefore, the projection equation is calculated. The projected radial distance based on the radial/angular world-to-camera projection is as followed:

$$\Rightarrow r_{\text{image}} = f(\varphi) = \frac{a_1 \sin(\varphi)}{a_2 + \cos(\varphi)} + V_{\text{res}}$$
Eq. 3.38

Assuming θ to be the planar angle in $X_s Y_s$ plane calculated from X_s axis with $R = 1$, we can calculate the projected coordinates in image plane:

$$X_s = \sin(\varphi) \cos(\theta), \quad Y_s = \sin(\varphi) \sin(\theta), \quad Z_s = \cos(\varphi)$$

$$x_p = r_{\text{image}} \cos(\theta) = a_1 \frac{X_s}{a_2 + Z_s} \quad y_p = r_{\text{image}} \sin(\theta) = a_1 \frac{Y_s}{a_2 + Z_s}$$
Eq. 3.39

Now if we fix the parameter a_1 by assuming $a_1 = f \cdot S_0 = 1$, we can rebuild a special projection case comparable to C. Mei model (Mei & Rives, 2007). However, in this effort we will continue with the more generalized projection transformation.

3.7.2. General Angular C2W Projection

In this chapter, we derive the inverse or C2W transformation of C. Mei model from the equations used in the previous chapter. As discussed in previous chapter, these simplifications help us to increase the calculation performance in runtime scenarios and simplify the formalism:

$$\begin{aligned}
 r_{\text{image}} = f(\varphi) &= \frac{a_1 \sin(\varphi)}{a_2 + \cos(\varphi)} \quad \Rightarrow \quad r a_2 + r \cos(\varphi) = a_1 \sqrt{1 - \cos^2(\varphi)} \\
 \Rightarrow \quad r^2 a_2^2 + r^2 \cos^2(\varphi) + 2r^2 a_2 \cos(\varphi) &= a_1^2 (1 - \cos^2(\varphi)) \\
 \Rightarrow \quad (r^2 + a_1^2) \cos^2(\varphi) + 2r^2 a_2 \cos(\varphi) + r^2 a_2^2 - a_1^2 &= 0
 \end{aligned} \tag{Eq. 3.40}$$

Now we estimate the parameters φ analytically to be able to calculate the sight-angle from the image coordinates:

$$\begin{aligned}
 \cos(\varphi) &= \frac{-r^2 a_2 + \sqrt{r^4 a_2^2 - (r^2 + a_1^2)(r^2 a_2^2 - a_1^2)}}{r^2 + a_1^2} \\
 \cos(\varphi) &= \frac{-r^2 a_2 + a_1 \sqrt{r^2(1 - a_2^2) + a_1^2}}{r^2 + a_1^2} \\
 \text{and} \quad \sin(\varphi) &= \frac{r a_1 a_2 + r \sqrt{r^2(1 - a_2^2) + a_1^2}}{r^2 + a_1^2}
 \end{aligned} \tag{Eq. 3.41}$$

We continue by defining the inverse transformation projection factors C_1, C_2 and C_3 and calculating the C2W projection angle based on the radial distance in image plane:

$$\begin{aligned}
 C_1 &= -\frac{a_2}{a_1^2} = -\frac{f + H_v}{f^2 S^2} \quad \text{and} \quad C_2 = \frac{1 - a_2^2}{a_1^2} = \frac{1 - (f + H_v)^2}{f^2 S^2} \quad \text{and} \quad C_3 = \frac{1}{a_1^2} = \frac{1}{f^2 S^2} \\
 \cos(\varphi) &= \frac{C_1 r^2 + \sqrt{C_2 r^2 + 1}}{C_3 r^2 + 1} \quad \Rightarrow \quad \varphi = \arccos\left(\frac{C_1 r^2 + \sqrt{C_2 r^2 + 1}}{C_3 r^2 + 1}\right) \quad \text{and} \quad C_1, C_2, C_3 = \text{cte.}
 \end{aligned} \tag{Eq. 3.42}$$

We can now use the sight angle φ in the C2W calculations. In the next step, we derive the equations used for calculating direction vectors (3D sphere coordinates) from image coordinates:

$$\begin{aligned}
 Z_s &= R \cos(\varphi) = R \frac{C_1 r^2 + \sqrt{C_2 r^2 + 1}}{C_3 r^2 + 1}, \quad R_{XY} = R \sin(\varphi) = \sqrt{R^2 - Z_s^2} \\
 X_s &= R_{XY} \cos(\theta) = R \sin(\varphi) \frac{x_p}{r}, \quad Y_s = R_{XY} \sin(\theta) = R \sin(\varphi) \frac{y_p}{r} \\
 \Rightarrow \quad \begin{cases} Z_s = R \frac{C_1 r^2 + \sqrt{C_2 r^2 + 1}}{C_3 r^2 + 1} \\ X_s = \frac{\bar{x}_p}{r} \sqrt{R^2 - Z_s^2} \\ Y_s = \frac{\bar{y}_p}{r} \sqrt{R^2 - Z_s^2} \end{cases} & \quad @ \quad C_1..C_3 = \text{CTE.}
 \end{aligned} \tag{Eq. 3.43}$$

Using Eq. 3.43, we can calculate the direction vectors directly from the image coordinates using the three coefficients C_1, C_2, C_3 . We can also calculate the inverse transformation in other forms. The inverse of the scale factor would be:

$$\lambda = \frac{1}{S} = \frac{f + H_v + Z_s}{f} \quad \text{Eq. 3.44}$$

Therefore, we can write the simplified C2W equations in the following form:

$$\Rightarrow \begin{cases} Z_s = R \frac{C_1 r^2 + \sqrt{C_2 r^2 + 1}}{C_3 r^2 + 1} \\ X_s = R \bar{x}_p \frac{f + H_v + \frac{Z_s}{R}}{f} \\ Y_s = R \bar{y}_p \frac{f + H_v + \frac{Z_s}{R}}{f} \end{cases} \quad @ \quad C_1..C_3 = \text{CTE.} \quad \text{Eq. 3.45}$$

Fig. 3.5 illustrates examples of real-time image rectification using W2C (3D-2D) and C2W (2D-3D) transformation discussed in [3.7.1] and [3.7.2].



a	b	c
d	e	f

Fig. 3.5. Image rectification using the C2W (2D-3D) and W2C (3D-2D) transformations
a, b, c: original images from an omni-directional camera installed in vehicle
d, e, f: rectified images corresponding to the above images respectively
(Courtesy of Robert Bosch GmbH)

Eq. 3.45 describes a general camera-to-world transformation that gives us many possibilities for parameter modifications. Again, if we fix the parameter a_1 by assuming $a_1 = f \cdot S_0 = 1$, we can rebuild the original C. Mei projection transformation (Mei & Rives, 2007). This implies having unified normal image coordinates, and $R = 1$, which means projection sphere with unit radius. However, we will continue with the more generalized formalism derived in Eq. 3.45 to recommend certain runtime optimization.

3.7.3. Runtime optimization of C2W transformation

As we previously described the image-rectification of fisheye cameras in [3.7.2] and illustrated the image rectification in Fig. 3.5 using the C2W and W2C transformation, we see benefits and potential improvements from real-time optimization. We can rewrite the inverse transformation by assuming the projection factor C_2 to be equal zero. This helps us directly eliminate the square root term as a further simplification of the C2W inverse transformation, which helps us to optimize the real-time calculation:

$$\begin{aligned} C_2 = \frac{1 - (f + H_v)^2}{f^2 S^2} = 0 &\Rightarrow f + H_v = 1 \\ \Rightarrow -R_s \sin\left(\frac{\alpha}{2}\right) \tan\left(\frac{\alpha}{2}\right) - R_s \cos\left(\frac{\alpha}{2}\right) = 1 &\Rightarrow \cos\left(\frac{\alpha}{2}\right) = R_s \end{aligned} \quad \text{Eq. 3.46}$$

Eq. 3.46 implies that by defining the specifications of the inner virtual perspective projection, the projection sphere will have a radius equal to the cosine of the angle α . This will eventually mean that α can have a maximum value of 180 [deg]. Therefore this projection optimization is realizable inside a FOV of 180 [deg]. Now after evaluating α , we select $R_s = \cos\left(\frac{\alpha}{2}\right)$, $C_2 = 0$ and f, H_v . Based on this R_s the optimized C2W transformation will be rewritten.

$$\begin{cases} Z_s = R_s \frac{C_1 r^2 + 1}{C_3 r^2 + 1} \\ X_s = \bar{x}_p \frac{f + H_v + Z_s}{f} \\ Y_s = \bar{y}_p \frac{f + H_v + Z_s}{f} \end{cases} \quad \text{Eq. 3.47}$$

Eq. 3.47 is the further simplified C2W and W2C transformation. We discussed in [1.2] that we try to reduce the required computational performance of our functions on low-power devices to increase the calculation speed. One of the main advantages of this model described in Eq. 3.47 is that we do not have any square root operation in the above formalism. Extra operations such as square root calculation usually requires iterative or expansion series estimation that cost calculation time especially if we are performing the calculations per pixel. Eq. 3.47 avoids square root calculation on the processor and improves the real-time calculation performance. Using these optimizations to calculate values for camera-to-world and world-to-camera transformation makes it possible to perform a real-time image rectification easier on mobile hardware. The deviations from the optical projection could be compensated by the distortion function. It is necessary to mention, as discussed in Eq. 3.23, to reduce the correlation problem between f and H_v we can consider the parameter H_v as pre-defined fixed parameter and only estimate f . Alternatively, we can reparametrize $f + H_v$ as a single parameter like H_{v2} and optimize them consecutively to reduce correlation in iterative least squares estimation.

In this chapter, we developed an extended omnidirectional projection model based on the model from (Mei & Rives, 2007) that physically models the behaviour of optical module, and can be further optimized to run efficiently on low power or mobile processors without consuming huge processing powers. As discussed in chapter [3], one of the main expansion of this omnidirectional model is the addition of a flexible and efficient distortion model, i.e. the free-function model. In the next chapter, we introduce the concept of this model for both fisheye and perspective cameras, discuss its real-time recalculation and rectification characteristics, compare it with other distortion models, and discuss the details.

4. The Free-Function Calibration Distortion Model

As discussed in previous chapters [2.4], camera calibration is the evaluation of parameters and the specification of a world-to-camera projection. In most of the self-calibration solutions, we define the camera model and specifically the distortion model and try to find the model coefficients by solving the general least squares equations system. In simple words the calibration is a classical mathematical function fitting problem (Tang, et al., 2012 A) (Tang, et al., 2012 B).

The accuracy of the calibration depends on the accuracy of input data such as the image observations, the known three-dimensional objects and the most important the distortion model itself. The better the distortion model fits itself to the real image distortion - while still having the numerical stability - the higher the calibration accuracy while assuming accurate image measurements. We should mention the recent advancements in camera calibration (Baretto & Daniilidis, 2004) (Ying, et al., 2006) (Scaramuzza, et al., 2006) (Mei & Rives, 2007) (Scaramuzza, 2014) and introduction of new models which have specific advantages over the previous/old models (Tang, et al., 2012 A) (Tang, et al., 2012 B). Still many of the recent calibration models use physically-motivated distortion models similar to, or based on Brown (Brown, 1971) (Brown, 1976) (Clarke & Fryer, 1998).

Here we have tried to look at the above problem from another aspect of view. Instead of finding the best calibration model or mathematically the best function to fit to our optical distortion, we assume a hypothetical function in our calibration model or in other words the calibration model to be analytically unknown (Nekouei Shahraki & Haala, 2015 A). To be able to practically use an unknown function, we can adopt a generic function expansion with unknown coefficients as equivalent and evaluate those coefficients. By setting the functions to be unknown, we remove any analytical limitations and explicit constraint in the model. This implies that the functions would fit themselves to the observation data no matter what kind of mathematical form they possess. This increases the flexibility of the distortion model to fit itself to different lens distortions and furthermore be able to model the very local lens distortions; a subject which has been a concern for recent studies and researches (Luhmann, et al., 2011) and is difficult to achieve in conventional model.

The goals for developing this model is similar to the targets set by finite-element (FE) modelling (Lichti & Chapman, 1995) (Luhmann, 2010) (Reznicek & Luhmann, 2019) to tackle the problem of the remaining systematic errors in the camera projection, when e.g. considering physical models such as Brown-Fraser (Brown, 1971) (Fraser, 1997). These errors may be caused by unflatness in the projection plane or the imaging sensor, aberrations and refractions caused by various deviations in the optical elements from the design geometry and so on (Reznicek & Luhmann, 2019). Contrary to the finite element approach, in free-function distortion modelling we try to model most of these errors in the distortion modelling using distortion functions based on primary bivariate functions. The primary functions are then be approximated by bivariate polynomials (Weber, 2012) which then could be evaluated discretely utilizing least-squares estimation and using the observations in image domain. The results of the free-function model, comparable to the finite-element approach, could be represented as final distortion fields, which incorporate the systematic errors observed during camera calibration. Based on this model, the distortion modelling is not limited to the local distortion, but also modelling of the whole camera distortion field together with the local distortion is feasible.

Similar to the FE calibration method, to measure and model the systematic errors in the image we need to have the approximated three-dimensional coordinates of the points, the image coordinates and orientation of each image. In this work, in order to provide these data for the FF calibration, we perform the camera calibration once using a known camera model as the initial calibration (such as Brown, Tschebycheff, C. Mei, ...). It should be mentioned that because of the unknown form of the functions, there are difficulties using it directly in least-squares estimation systems. Therefore we need to use a conventional calibration model discussed in [§2.4] to model most of the distortion initially and then try to model the remaining local distortion in a second step as recommended by (Luhmann, et al., 2011) (Reznicek & Luhmann, 2019). This decision has the advantage that it provide us the necessary data and has the benefit of eliminating any possible correlations between the final distortion model and other variables such as extrinsic calibration parameters. This means the free-function modelling is performed as a posteriori calibration step after the initial calibration, and implies that extra correlation between calibration parameters will be non-existent.

We further develop the distortion model to be able to perform image undistortion. Undistortion is the operation to calculate the undistorted image coordinates from the distortion model or distorted coordinates (Mallon & Whelan, 2004). In real-time applications in driver-assistant systems and many other machine-vision systems, there is a need to have the undistortion model to perform direct camera-to-world (C2W) transformation. There are currently different ways to calculate the undistorted image points from the distorted points by calculating the inverse of the often radial-only distortion model using different numerical iterative solutions (Sebah & Gourdon, 2011) (Kreyszig, 2010) or analytical solutions (Mei & Rives, 2007) which are only available for simpler low-order radial distortions. Other solutions are for example using the same distortion model for inverse calculation of only-radial distortion (De Villiers, et al., 2008) which face the same real-time calculation challenges on low-TDP Platforms. The iterative solutions are basically time consuming and are not suitable for real-time applications. It should also be mentioned that to have higher accuracies, the number of iterations should increase which again slows the whole undistortion process. Thus a straightforward solution is needed to provide us with accuracies around 0.01-0.02 pixel which is much higher than current other solutions (Mallon & Whelan, 2004) (De Villiers, et al., 2008). This enable us to perform undistortion in real-time much faster. For this purpose, we have used the free-function model to calculate the undistortion model from the distortion model, which has the advantage and possibility of being performed - using only the distortion parameters - at a later stage after the first calibration.

As also discussed in chapter 1.2, we aim at optimizing our calibration solution for low-power processing hardware with limited processing resources and consider real-time recalculation scenarios. Therefore, in order to implement the free-function model on the hardware side, we will consider the capabilities of hardware processing to increase the real-time calculation performance, and compare it to other distortion models.

In the next chapter, we start to study the mathematical concept of the free-function model and continue with the details of its implementation and optimization.

4.1. The Mathematics of the Free-Function-Model

In this chapter, we study the mathematics behind the free-function calibration model, discuss its specifications and review the different ways we can take advantage of this model. First, we write the general self-calibration equation:

$$\begin{aligned}x_{\text{bdist}} &= x_{\text{bundist}} + \Delta x_{x_{\text{b}}, y_{\text{b}} \text{undist}} \\y_{\text{bdist}} &= y_{\text{bundist}} + \Delta y_{x_{\text{b}}, y_{\text{b}} \text{undist}}\end{aligned}\tag{Eq. 4.1}$$

In Eq. 4.1, $\Delta x_{x_{\text{b}}, y_{\text{b}} \text{undist}}$ and $\Delta y_{x_{\text{b}}, y_{\text{b}} \text{undist}}$ are functions of x_{bundist} and y_{bundist} (or x_{b} and y_{b}) and describe the distortion per pixel. Now considering $\Delta x_{x_{\text{b}}, y_{\text{b}} \text{undist}}$ and $\Delta y_{x_{\text{b}}, y_{\text{b}} \text{undist}}$ the free-function distortion model is defined as multiplication of two bivariate functions:

$$\begin{aligned}\Delta x_{x_{\text{b}}, y_{\text{b}} \text{undist}} &= \sum_{i=1}^n f_i(x_{\text{b}}, y_{\text{b}}) \cdot k_i(x_{\text{b}}, y_{\text{b}}) \\ \Delta y_{x_{\text{b}}, y_{\text{b}} \text{undist}} &= \sum_{i=1}^n g_i(x_{\text{b}}, y_{\text{b}}) \cdot l_i(x_{\text{b}}, y_{\text{b}})\end{aligned}\tag{Eq. 4.2}$$

The expansion of the above equation will be as follows:

$$\begin{aligned}\Delta x_{x_{\text{b}}, y_{\text{b}} \text{undist}} &= f_1(x_{\text{b}}, y_{\text{b}}) \cdot k_1(x_{\text{b}}, y_{\text{b}}) + f_2(x_{\text{b}}, y_{\text{b}}) \cdot k_2(x_{\text{b}}, y_{\text{b}}) + f_3(x_{\text{b}}, y_{\text{b}}) \cdot k_3(x_{\text{b}}, y_{\text{b}}) + \dots \\ \Delta y_{x_{\text{b}}, y_{\text{b}} \text{undist}} &= g_1(x_{\text{b}}, y_{\text{b}}) \cdot l_1(x_{\text{b}}, y_{\text{b}}) + g_2(x_{\text{b}}, y_{\text{b}}) \cdot l_2(x_{\text{b}}, y_{\text{b}}) + g_3(x_{\text{b}}, y_{\text{b}}) \cdot l_3(x_{\text{b}}, y_{\text{b}}) + \dots\end{aligned}\tag{Eq. 4.3}$$

In Eq. 4.2, f_i , k_i , g_i and l_i are functions of $(x_{\text{b}}, y_{\text{b}})$. These functions will later be further expanded, whose expansion coefficients have to be evaluated numerically. This is further discussed in [4.1.4] and [4.2].

4.1.1. Function Superposition

Having Eq. 4.2, we can see that the functions k_i and l_i could be selected arbitrarily as long as the functions f_i and g_i remain unknown. In other words, by holding f_i and g_i functions as unknown (free) we can still model the image distortion using arbitrary forms that we put on k_i and l_i :

$$f_i(x_b, y_b) \cdot k_i(x_b, y_b) = f'_i(x_b, y_b) \cdot k'_i(x_b, y_b) \quad \text{Eq. 4.4}$$

and:

$$f_i(x_b, y_b) \cdot k_i(x_b, y_b) \cdot k_i^{-1}(x_b, y_b) = f'_i(x_b, y_b)$$

In Eq. 4.4 the function k'_i is the arbitrary form of k_i function and f'_i is the respective f_i function after superposing k'_i . This feature is most important because we can enforce specific scaling between f_i and k_i (the same for g_i and l_i). It could otherwise in some cases introduce singularities if f_i and k_i are to some extents numerically correlated at the initialization phase of LSE. This helps us partially enforce the numerical modelling of image distortion up to very certain exponents, which should comply with the available hardware processing capabilities for different applications such as real-time processing.

Assuming known function forms for k'_i and l'_i in the original expansion we have the free-function model with superposed functions:

$$\begin{aligned} \Delta x_{x_b, y_b \text{ undist}} &= \sum_{i=1}^n f'_i(x_b, y_b) \cdot k'_i(x_b, y_b) \\ \Delta y_{x_b, y_b \text{ undist}} &= \sum_{i=1}^n g'_i(x_b, y_b) \cdot l'_i(x_b, y_b) \end{aligned} \quad \text{Eq. 4.5}$$

Comparing Eq. 4.2 with Eq. 4.5, we can see that arbitrary functions could be used inside the free-function calibration model. In case we select the arbitrary forms for k' as functions of x_b and l' as a functions of y_b then the functions f and g could therefore be evaluated discretely on each image column and row:

$$\begin{aligned} \Delta x_{x_{b1}, y_b \text{ undist}} &= f'_1(x_{b1}, y_b) \cdot k'_{1}(x_{b1}) + f'_2(x_{b1}, y_b) \cdot k'_{2}(x_{b1}) + \dots & @ x_b = x_{b1}, y_b = y_b \\ \Delta y_{x_b, y_{b1} \text{ undist}} &= g'_1(x_b, y_{b1}) \cdot l'_{1}(y_{b1}) + g'_2(x_b, y_{b1}) \cdot l'_{2}(y_{b1}) + \dots & @ x_b = x_b, y_b = y_{b1} \end{aligned} \quad \text{Eq. 4.6}$$

We can take advantage of this feature especially for developing calculation strategies and run-time optimizations, which would be later discussed in [4.2].

4.1.2. Free-Function Model for Distortion Estimation

Based on the super positioning concept from [4.1.1], we can for example define the functions k' and l' to be $k'_i(x_b, y_b) = y_b^{ij}$ and $l'_i(x_b, y_b) = x_b^{ij}$. Therefore, we only need to estimate f_i and g_i functions (respectively f'_i and g'_i) using the numerical estimation given in [4.1.4]. Therefore, using the original function definition we have:

$$\Delta x_{x_b, y_b \text{ undist}} = \sum_{i=1}^n f_i(x_b, y_b) \cdot y_b^{ij} = \sum_{i=1}^n f_i(x_b, y_b) \cdot y_b^i \quad @ j = 1$$

$$\Delta y_{x_b, y_b \text{ undist}} = \sum_{i=1}^n g_i(x_b, y_b) \cdot x_b^{ij} = \sum_{i=1}^n g_i(x_b, y_b) \cdot x_b^i \quad @ j = 1$$
Eq. 4.7

The functions can therefore be evaluated discretely on each image column/row. The discrete estimation of free-function model using arbitrary functions is given in Eq. 4.8:

$$\Delta x_{x_{b1}, y_b \text{ undist}} = f_1(x_{b1}, y_b) \cdot y_b + f_2(x_{b1}, y_b) \cdot y_b^2 + \dots \quad @ x_b = x_{b1}, y_b = y_b$$

$$\Delta y_{x_b, y_{b1} \text{ undist}} = g_1(x_b, y_{b1}) \cdot x_b + g_2(x_b, y_{b1}) \cdot x_b^2 + \dots \quad @ x_b = x_b, y_b = y_{b1}$$
Eq. 4.8

Having the $\Delta x_{x_{b1}, y_b}$ and $\Delta y_{x_b, y_{b1}}$ as image observation sets, by performing the LSE on each column or row we can numerically estimate f_i and g_i functions. The accuracy of fitting depends on the degree of the function series which is indicated by n .

Generally, by using higher degrees (higher length for the function series) we can have higher modelling accuracy. The higher the degree n , the higher is also the amount of parameters that needed to be stored as calibration data in memory, and the higher amount of processing power needed to recalculate the distortion values per pixel. Therefore, the minimum degree is selected based on the required accuracy threshold, to make sure that the accuracy requirement is fulfilled, and in the same time the minimum amount of memory is allocated to store the calibration data and the least amount of processing is needed for achieving real-time distortion recalculation

4.1.3. Free-Function Model for Undistortion Estimation

The inverse model of distortion together with the distortion model are used to be able to, for example, calculate the three dimensional direction vectors directly from input data. Generally, for undistortion modelling we can write:

$$x_{b\text{undist}} = x_{b\text{dist}} + \Delta x_{x_b, y_b \text{dist}} \quad y_{b\text{undist}} = y_{b\text{dist}} + \Delta y_{x_b, y_b \text{dist}} \quad \text{Eq. 4.9}$$

In Eq. 4.9 the $\Delta x_{x_b, y_b \text{undist}}$ and $\Delta y_{x_b, y_b \text{undist}}$ are functions of $x_{b\text{dist}}$ and $y_{b\text{dist}}$ (or x_{bd} and y_{bd}). Now the free-function model is defined:

$$\begin{aligned} \Delta x_{x_b, y_b \text{dist}} &= \sum_{i=1}^n f_i(x_{bd}, y_{bd}) \cdot k_i(x_{bd}, y_{bd}) \\ \Delta y_{x_b, y_b \text{dist}} &= \sum_{i=1}^n g_i(x_{bd}, y_{bd}) \cdot l_i(x_{bd}, y_{bd}) \end{aligned} \quad \text{Eq. 4.10}$$

Assuming known or arbitrary function forms for k'_i and l'_i in the original expansion, for undistortion calculation, we would have:

$$\Delta x_{x_b, y_b \text{dist}} = \sum_{i=1}^n f_i(x_{bd}, y_{bd}) \cdot k'_i(x_{bd}, y_{bd}) \quad \text{Eq. 4.11}$$

The functions could be selected as $k'_i(x_{bd}, y_{bd}) = y_{bd}^{ij}$ and $l'_i(x_{bd}, y_{bd}) = x_{bd}^{ij}$ in the original expansion. Therefore we would have:

$$\Delta x_{x_b, y_b \text{dist}} = \sum_{i=1}^n f_i(x_{bd}, y_{bd}) \cdot y_{bd}^{ij} = \sum_{i=1}^n f_i(x_{bd}, y_{bd}) \cdot y_{bd}^i \quad @ j = 1 \quad \text{Eq. 4.12}$$

and

$$\Delta y_{x_b, y_b \text{dist}} = \sum_{i=1}^n g_i(x_{bd}, y_{bd}) \cdot x_{bd}^{ij} = \sum_{i=1}^n g_i(x_{bd}, y_{bd}) \cdot x_{bd}^i \quad @ j = 1$$

Similar to distortion estimation, in order to estimate the free-function model for undistortion we would need an initial calibration e.g. with a physical or mathematical distortion model first.

4.1.4. Discrete Expansion Equivalency

In order to evaluate the functions f' and g' numerically and discretely on each image pixel x and y , a discrete numerical estimation model from (Weber, 2012) is used:

$$f(u, v) := \begin{pmatrix} f^x(u, v) \\ f^y(u, v) \end{pmatrix} \approx \begin{pmatrix} \sum_{k=1}^n f_k^x(u) \cdot h_k^x(v) \\ \sum_{k=1}^n g_k^y(u) \cdot h_k^y(v) \end{pmatrix} \quad \text{Eq. 4.13}$$

The Eq. 4.13 presumes that a function based on a multivariate parameterisation such as f^x and f^y could be separated into a product of bivariate or univariate analytical functions which are noted as f_k^x, h_k^x, g_k^y and h_k^y . The equivalent expansion of these functions could then be used for numerical approximation. The functions h_k^y and h_k^x are bivariate or two univariate functions, described as polynomials with unknown coefficients. The parameter n being the degree of expansion and regarded as the degree of the free-function model. This means that the currently unknown functions f_k^x, h_k^x, g_k^y and h_k^y will be estimated numerically.

Eq. 4.13 could be used to describe the functions of the intrinsic model as discrete functions in the form of series with u and v parameters. Using this discrete estimation technique in Eq. 4.13 in free-function model (Eq. 4.5 or Eq. 4.10), the functions would expand (Weber, 2012) with LH_k and LV_k being the discrete estimations of f_k, g_k and h_k independently which could then be calculated numerically:

$$\begin{aligned} \Delta x_{x_b, y_b \text{ undist}} &= \sum_{id=1}^n LH_{id}^f(i) \cdot LV_{id}^{k'}(j) = LH_1^f(i) \cdot LV_1^{k'}(j) + LH_2^f(i) \cdot LV_2^{k'}(j) + \dots \\ \Delta y_{x_b, y_b \text{ undist}} &= \sum_{id=1}^n LH_{id}^g(i) \cdot LV_{id}^{l'}(j) = LH_1^g(i) \cdot LV_1^{l'}(j) + LH_2^g(i) \cdot LV_2^{l'}(j) + \dots \end{aligned} \quad \text{Eq. 4.14}$$

Discrete numerical estimation of free-function model is given in Eq. 4.14. By considering the discrete form of the free-function equations and performing function evaluation (by means of least squares estimation or other estimation methods) on each image column/row we can numerically estimate f_i and g_i functions. Based on the discussions in previous chapter we know the f_i and g_i functions or respectively f'_i and g'_i and their equivalent expansions are based on (x_b, y_b) . This means an initial distortion value on pixel level is required in order to use LSE. This would be satisfied by initially using a standard distortion model in the calibration. After estimating the initial distortion, we can calculate the initial distortion value for each pixel and use as input for the free-function distortion estimation. The flexibility of this model allows us to further consider the local distortions in the modelling. This is further discussed in [4.3].

4.2. Implementation, Real-Time Calculation and Optimization

As discussed at the beginning of chapter [4] in the introduction of the free-function model and also one of the motivations of this work, we try to optimize the distortion model for real-time scenarios where the distortion is recalculated and the images are rectified in real-time. Any calculation generally takes some amount of time on any processor, but certain operations are more desired, from which we can mention vector multiplication. Generally, multiplication of $[m, n]$ and $[n, m]$ vectors could easily be done in real-time in variety of hardware based on FPGA, x86 and ARM processors and general-purpose GPUs without much of a performance hit. Therefore this criterion is considered in the implementation and optimization.

We optimize the free-function model for optimal implementation in the hardware side. These optimizations help us to have an efficient use of the available memory and maximize the computing performance which is a key factor for fast run-time processing as discussed above. As the direct result of Eq. 4.5 as well as Eq. 4.7 we can set the free functions $k'_i = l'_i = A_i(x_b, y_b)$. This would imply using the same arbitrary functions A_i as the known part for function superposition to model the distortion in both x and y directions:

$$\Delta x_{x_b, y_b \text{ undist}} = \sum_{i=1}^n f_i(x_b, y_b) \cdot A_i(x_b, y_b) \quad \Delta y_{x_b, y_b \text{ undist}} = \sum_{i=1}^n g_i(x_b, y_b) \cdot A_i(x_b, y_b) \quad \text{Eq. 4.15}$$

The function $A_i(x_b, y_b)$ is calculated once in order to estimate both $f_i(x_b, y_b)$ and $g_i(x_b, y_b)$. After we estimate f_i and g_i functions, the discretely evaluated functions could be saved in device memory and later be used in corresponding applications. The distortion calculation process is then performed according to Eq. 4.14 as f_i and g_i are available for each image column/row.

As an optimization solution, in case of selecting the arbitrary functions A_i as functions of y_b coordinates only i.e. $A_i(y_b)$, both f_i and g_i functions would be evaluated for each image column. In this case a significant portion of required memory (to save the discrete evaluated functions) would be saved:

$$\Delta x_{x_b, y_b \text{ undist}} = \sum_{i=1}^n f_i(x_b, y_b) \cdot A_i(y_b) \quad \Delta y_{x_b, y_b \text{ undist}} = \sum_{i=1}^n g_i(x_b, y_b) \cdot A_i(y_b) \quad \text{Eq. 4.16}$$

Therefore, using the discrete expansion in Eq. 4.14, the distortion model looks like:

$$\Delta x_{x_b, y_b \text{ undist}} = \sum_{i=1}^n LH_i^f(x_b) \cdot LV_n^A(y_b) \quad \Delta y_{x_b, y_b \text{ undist}} = \sum_{i=1}^n LH_i^g(x_b) \cdot LV_n^A(y_b) \quad \text{Eq. 4.17}$$

Eq. 4.17 provides the optimization for undistortion modelling using superposition on arbitrary functions. This means that the expansion parameters of $LH_i^f(x_b)$ and $LH_i^g(x_b)$ should be evaluated numerically on all of the image columns with the parameter n being the degree of the expansion.

The functions $LH_i^f(x_b)$ and $LH_i^g(x_b)$ and LV_i^A could be conveniently represented in discrete form alongside the image row and columns as they are discretely estimated with the parameter count equivalent or proportional to the dimensions or domain of the observations which both correspond to pixel count in two dimensions.

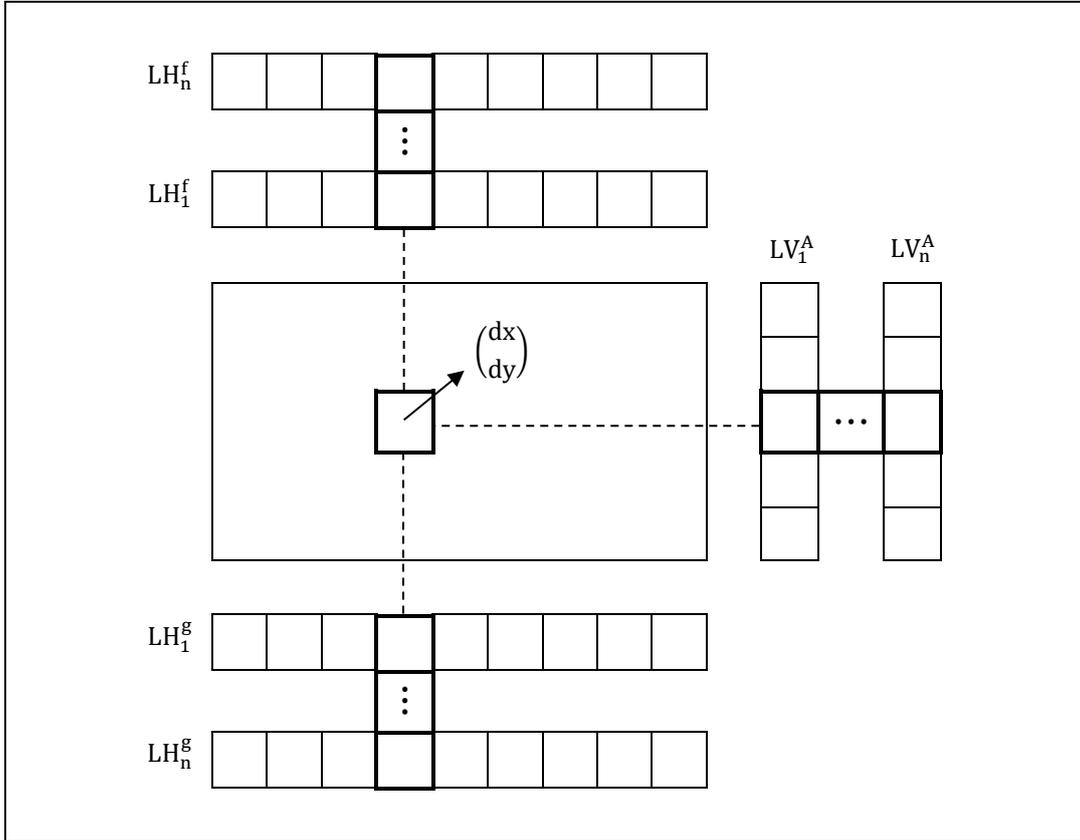


Fig. 4.1 Structure of the calibration model (Optimization 1)

The structure of the estimated intrinsic model is illustrated in Fig. 4.1. This illustration is motivated by, and based on the illustration used by (Weber, 2012) as it differentiates between vector multiplications in different directions. The numerical estimation of $f_i(x_b, y_b)$ and $g_i(x_b, y_b)$ is represented by LH_i^f and LH_i^g while the numerical form of $A_i(y_b)$ is represented as LV_i^A in form of LUTs (Look-Up Tables).

As an alternative optimization solution, in case we select the arbitrary function A as a function of x_b coordinates only, i.e. $A_i(x_b)$, both $f_i(x_b, y_b)$ and $g_i(x_b, y_b)$ functions would be evaluated for each column. In this case too, a significant portion of required memory used to save the discrete evaluated functions would be saved.

$$\Delta x_{x_b, y_b \text{ undist}} = \sum_{i=1}^n f_i(x_b, y_b) \cdot A_i(x_b) \quad \Delta y_{x_b, y_b \text{ undist}} = \sum_{i=1}^n g_i(x_b, y_b) \cdot A_i(x_b) \quad \text{Eq. 4.18}$$

And using the discrete expansion in Eq. 4.14, the distortion model looks like:

$$\Delta x_{x_b, y_b \text{ undist}} = \sum_{i=1}^n LV_i^f(y_b) \cdot LH_n^A(x_b) \quad \Delta y_{x_b, y_b \text{ undist}} = \sum_{i=1}^n LV_i^g(y_b) \cdot LH_n^A(x_b) \quad \text{Eq. 4.19}$$

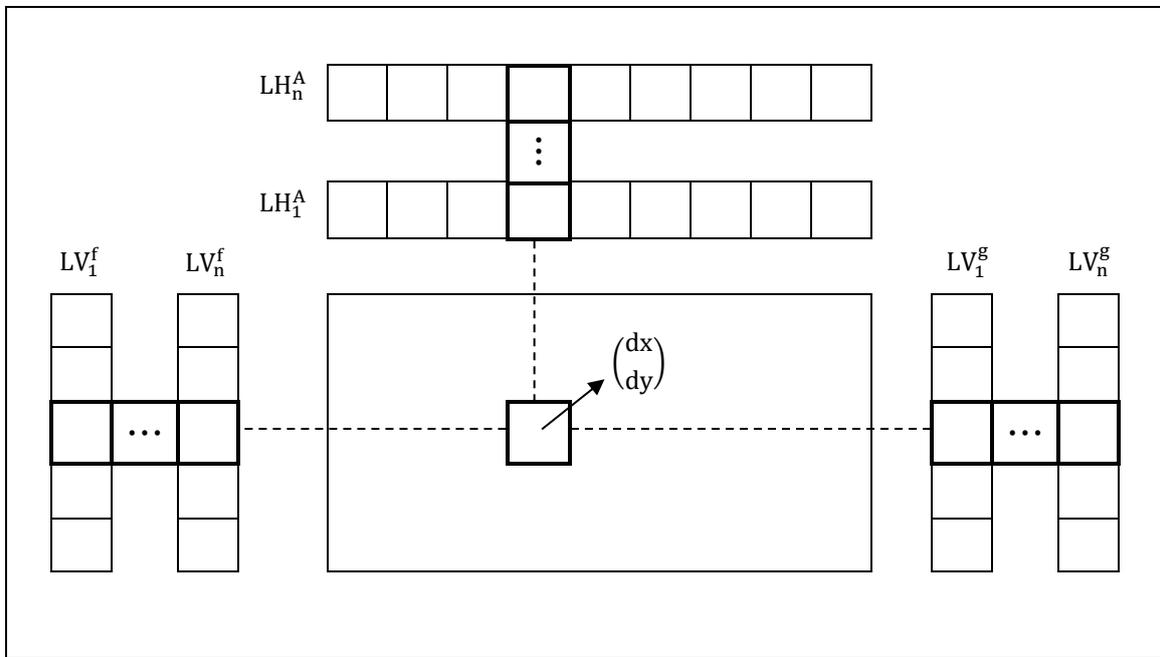


Fig. 4.2 Structure of the calibration model, Optimization 2

This representation in in Fig. 4.2 is motivated by (Weber, 2012) and helps us get a better understanding of the interpretation of the parameters. The numerical estimation of $f_i(x_b, y_b)$ and $g_i(x_b, y_b)$ is represented by LV_i^f and LV_i^g while the numerical form of $A_i(x_b)$ is represented as LH_i^A in form of LUTs (Look-Up Tables).

4.2.1. Performance Evaluation

As we can see in Fig. 4.1 and Eq. 4.16 and alternatively in Fig. 4.2 and Eq. 4.19 there is only a linear multiplication operation per pixel necessary in order to calculate the distortion values in x and y directions. This drastically improves the computation performance in run-time applications (Nekouei Shahraki & Haala, 2015 A). The above is an important feature that significantly reduces the demand for high computational power in our hardware platforms and therefore makes the run-time calculation process performable on low TDP processing units. As an effort to achieve a measure of performance, similar to the calculation of floating-point operation per second for processors (FLOP), we can approximately calculate the floating-point operation per pixel for different models. We calculate the sum of addition and multiplications and compare them together to get a sense of the theoretical real-time performances.

Distortion Model	CPU operations per point per pixel
Brown-Fraser	≈ 60
Brown Mixed-AP	≈ 100
Tschebycheff Deg. 3	> 100
Legendre Deg. (4,3)	> 100
Fourier Deg. 1	$\gg 100$
FF-Model Deg. 8	≈ 16
FF-Model Deg. 16	≈ 32

Table 4.1. Comparison of approximate required processing steps of the free-function model to different distortion models

As shown in Table 4.1, the free-function model (FF-Model) has a much less required processing power when calculated in real-time. Most of the distortion models that deliver high modelling accuracy require much more calculations per pixel. For example, the Fourier distortion model is based on trigonometric sub-functions as discussed in 2.4.3.9. These sub-functions are each being calculated using Maclaurin expansion (Kreyszig, 2010) with an expansion degree based on the function approximation accuracy. Therefore, these models require much more calculations.

Furthermore, one important advantage of the structure of this model is the simplicity of dividing the real-time calculation process i.e. linear multiplication operations to multiple smaller parts. This enables us to benefit from the performance enhancements of parallel processing on multi-core processing platforms or using GPGPU techniques.

The amount of allocated device memory needed to perform the image rectification, which is used to hold the information of the distortion model and calculate the distortion values for the image is very important. The amount of required memory depends on the image size and the complexity of the model (model degree n). Based on the optimization method and Fig. 4.1, the required memory equals to:

$$\text{Memory usage percentage} = 100 \cdot \frac{(2 \cdot \text{Width} + \text{Height}) \cdot n}{2 \cdot \text{Width} \cdot \text{Height}} \% \quad \text{Eq. 4.20}$$

Based on this optimization for an image size of 1024x512 for $n=4$ the required memory is around 0.97 % and for an image size of 1280x960 for $n=8$ the required memory is around 1.15 %. This corresponds to approximately only 81 KB of device memory and only 112 KB of run-time memory if a 32-Bit data structure is used. As visible in Fig. 4.1 and Fig. 4.2, it is also possible to change the resolution of the free-function to model the image distortion with desired step/resolution instead of the sensor native resolution which further contributes to the speed and to save the required device memory.

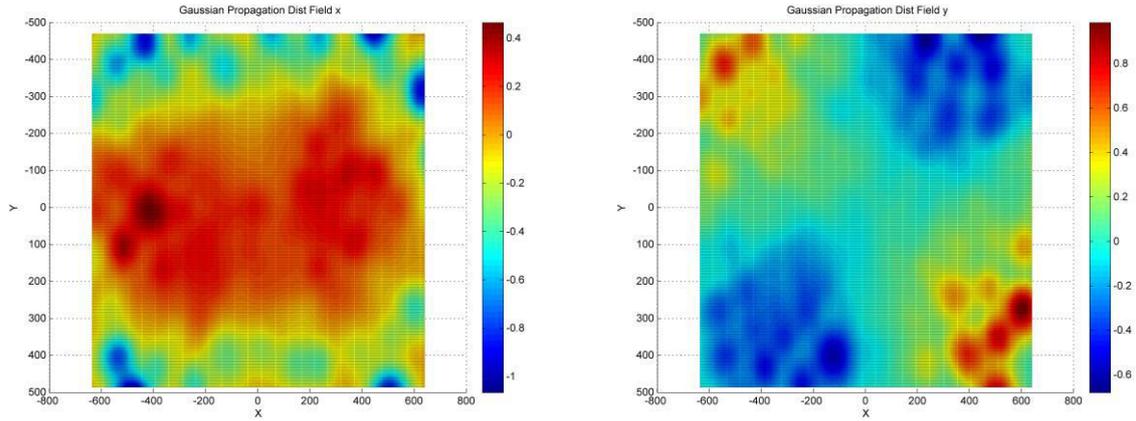
4.3. Modelling the Local Lens Distortions

To have accuracies beyond the initial distortion model - and most of the conventional calibration models - and be able to model the local distortions, the (initially) estimated distortion vector field should be updated using the remaining residuals on observed image points, which represent local lens distortions. Now we proceed to update the distortion-field for the local distortions. The cardinality condition is not present (Parasad & Lyengar, 1997) (Kirkwood, 2002) (Hummel, 2000) as we have fewer observations than the total pixel counts. The residual values on non-cardinal and non-integer pixel coordinates could be used as a second observation set to update the current distortion-map (per pixel). This is performed using the families of two-dimensional weighted interpolations. We can generally assume that the effects of local distortions would very gently decrease by increments in pixel distance from observation points. Therefore, in order to update the distortion-map, we can attach the local distortion data to the distortion dataset and use a - very smooth and continuous - interpolation method to interpolate the values for all the image points. As discussed in [2.1], we see that the behaviour of optical distortion is a continuous and smooth function. The interpolation requirement could be interpreted as a continuity criterion in the two-dimensional surface-derivative of the interpolated data. Therefore there are variety of interpolations that may not be suitable in this case. For example, the weighted IDW interpolation (Shephard, 1968) is very sensitive to the distance, and we do not get a homogenous distribution. Instead we will get sharp peaks in the interpolation space which practically do not resemble any real optical distortion.

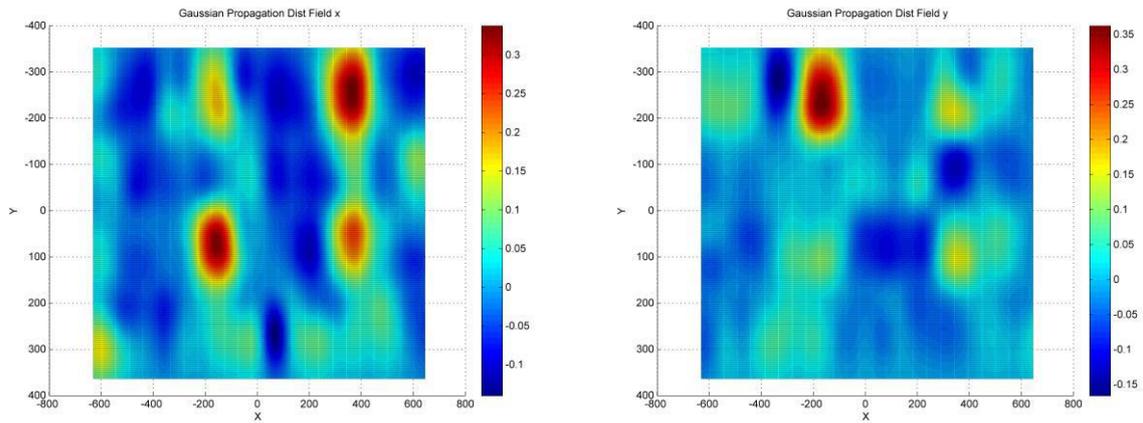
As discussed in this chapter, the optical behaviour of the camera can create local lens-distortion whose effects gradually increases in - and are limited to - some image areas. Then we discussed that the interpolation solution to the problem of constructing the local lens distortion field from the measured residual values in the image should comply with the derivative continuum criterion. Therefore we proceed by considering these local distortion as pulse responses of a set of two-dimensional Gaussian distribution functions. This is further backed-up by the previous studies on modelling the lens distortion using Gaussian functions (Ranganathan & Olson, 2012) and describing the optical effects such as aberration using Gaussian point-spread functions (Mahajan, 1998). This means that a Gaussian distribution function represents the estimated local lens distortion in an observation point in image.

$$G_{\text{Imp}}(x, y) = A e^{-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}} \quad \text{Eq. 4.21}$$

Eq. 4.21 is the two-dimensional Gaussian pulse function with pulse amplitude of A, standard deviations of σ_x, σ_y and centre position of x_0, y_0 in both x and y directions. As it is clear from the exponential formalism in Eq. 4.21, the two-dimensional Gaussian propagation conforms to the derivative continuum criterion so that the local distortion drops gradually and continuously by increasing the distance from the observation point, and vanishes afterward. To calculate the local distortion field, for each measured residual a pulse-response of a 2D Gaussian distribution function G_{Imp} is calculated on non-integer and non-cardinal pixel coordinates and added to a layer of cumulative distortion field using two-dimensional sub-pixel interpolation. After that, the results for each point would be calculated from the cumulative distortion field using weight values taken from the error distribution function to form a final distortion field. This is considered as the final local lens distortion-field and could be later used to update the originally estimated distortion field (Nekouei Shahraki & Haala, 2015 A).



a **b** Fig. 4.3. Example of estimated local distortion field (using Gaussian propagation) for a fisheye camera with high image local distortion. a: in image x direction, b: in image y direction
 Distribution domain size= 300, standard deviation factor = $\frac{1}{5}$. The colour-bar and axes coordinates are in pixel units. Courtesy of Robert Bosch GmbH



a **b** Fig. 4.4. Example of estimated local distortion field (using Gaussian propagation) for a perspective camera with high image local distortion. a: in image x direction b: in image y direction
 Distribution domain size= 300, standard deviation factor = $\frac{1}{5}$. The colour-bar and axes coordinates are in pixel units. Courtesy of Robert Bosch GmbH

Fig. 4.3 and Fig. 4.4 show the amount of the local lens distortion in both x, y directions for a fisheye and a central-projection camera. Having these data as input and adding them to the distortion-field generated by the original calibration model, we can perform the free-function model estimation. After that, by subtracting the generated distortion-field from the original model, we can evaluate how good the local distortions are modelled. We should mention that although the distortion field could be updated using the local lens distortion calculated from the residuals of the initial distortion model, the accuracy of the estimated local distortions and the final distortion field are limited to the accuracy of image observation and the three-dimensional target accuracy.

4.4. Calibration Strategy and Process Flow

In this chapter, we plan the necessary steps in order to calibrate our cameras in a systematic trend. These steps will be followed precisely at the calibration phase for each camera, and will be also considered when using the cameras for real-time image rectification. By considering process optimization, we can increase the speed of the calibration process-flow and also make sure of having a unified process for each camera. The latter is a necessity while calibrating a huge number of cameras one after each other in a series production. Because of that, we can ensure the same condition and quality for each device. This is important because we can hold the assumption that certain necessary conditions are valid for all of the cameras that are going to be used. The details of these steps are given in the following chapters.

4.4.1. Distortion Model and Initial Calibration

As we have also discussed in [4.1] and [4.1.2] we need to initially calibrate our cameras using a standard calibration model. It is because the calibration model that we introduced in [4] requires initial information about the camera IOR. It implies that we have two different steps that we use in order to calibrate the camera accurately, as also recommended by (Luhmann, et al., 2011) (Reznicek & Luhmann, 2019).

We can chose from variety of different models as a first attempt to understand the initial distortion model and the network geometry. In [2.4.3] we have introduced these projection models and calibration models that we use as a first-step calibration. This step is illustrated in Fig. 4.5, and is assigned with number 1.

4.4.2. Updating the Distortion Field and Local Distortion

As discussed in the previous chapter, as a necessary step in the calibration, we should provide the initial information before we start the final calibration process. The final calibration model is the combination of two parts: first, a camera projection model which fit to the special mapping designs of camera lenses, and second a distortion model which delivers the required accuracy. We need to extract and process the information related to the local distortion and feed them as complementary data to the final calibration model. This step is illustrated in Fig. 4.5, and is assigned with number 102 and 103.

4.4.3. Using Free-Function Calibration Model

This process is assigned with number 4 and 5 as illustrated in Fig. 4.5. As described in [4.1.2], [4.1.3] and [4.2] we perform free-function model using the previously prepared data from steps 1, 2 and 3 as illustrated in Fig. 4.5. The results would be given in the form of LUTs which are the discrete estimation of the optical distortion.

4.4.4. Finalizing the Calibration Process

As the final step, we would save the estimated parameters of the free-function model in the device memory. These parameters would be then recalled when using the camera to perform real-time image distortion rectification or other applications such as three-dimensional object reconstruction. This process is assigned with number 6 and 7 in Fig. 4.5

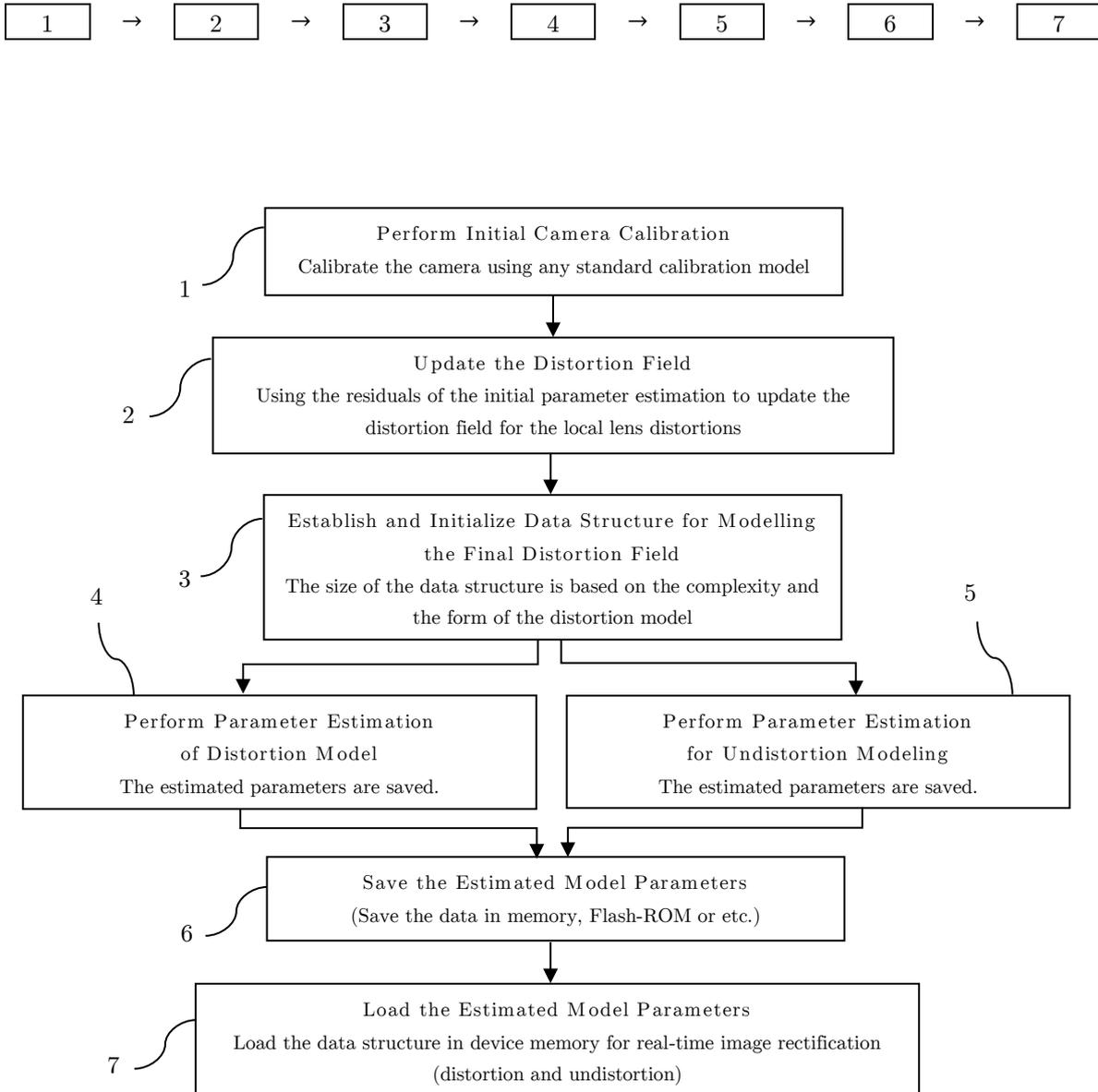


Fig. 4.5. Design of the calibration process flow

5. General Requirements for Calibration Test Stands

As discussed earlier [2.5.3], designing and developing optical-testing and calibration systems for performing various optical tests on every single camera is a requirement in computer-vision applications such as the video-based driver-assistant systems. Calibration test stands are often used when speed is the key requirement in performing camera calibration in a single-shot operation. However, there are also ever growing requirements for high calibration repetitive accuracy and high sub-pixel accuracy. For example, to calibrate the cameras accurately in a series production, we have to be able to perform single-shot calibration in a specific time limit (e.g. 5 second) for each camera. Therefore, a very high repetitive accuracy is required for the new calibration test stands. Based on the above criteria, we have designed calibration test stands, which would comply with those requirements and are further discussed in details.

5.1.1. Maintaining Calibration Accuracy for Complete Image Area

The calibration sub-pixel accuracy at the edges of image have a special importance in many applications in surround-view systems taking advantage of omnidirectional cameras. Also in Front-View driver assistant systems, these image areas include critical information about the movement of pedestrians, bicycle-riders and other vehicles in front of the vehicle. Therefore, we need to make sure that the designed test stand and the used calibration model will provide us with this accuracy at the edges.

When talking about the calibration accuracy at the edges, one of the most important criterions in this matter is the Runge's Phenomenon (Dahlquist & Björck, 2003). This phenomenon has been discussed in [2.4.3.10]. Therefore by considering the point-distribution and purposely designing it to be similar to the bundle-block point distribution, we can make sure that this phenomenon does not happen when performing the test stand calibration. This is also valid when using calibration models such as C. Mei model (Mei & Rives, 2007) or the free-function model [4] (Nekouei Shahraki & Haala, 2015 A) for which different types of polynomials or function series with degrees from 8 up to 32 are used for distortion modelling.

5.1.2. Additional Demands on Optical Tests

It is required and advantageous to have multiple functionalities available in our test systems. Therefore the test stand needs to be designed to encompass the capability for other optical tests on our cameras such as Sharpness-Testing (Estribeau & Magnan, 2004) or Stray-Light testing (Raizner, 2012). These optical tests provide us with further quality measures of our optical module such as the appearance of optical artefacts or unwanted optical effects such as ghosts in image. This helps us determine if a certain level of quality is reached for each camera. Therefore the existence of some of these extra functionalities should be considered in the test stand design.

When performing the camera intrinsic calibration, we also require the information about the alignment of the camera internal parts [2.5.3]. This information include the alignment of sensor with respect to the optical axis, the camera housing, or the reference camera coordinates system. Therefore we need to consider the capability of performing camera alignment testing in the design of test stand.

5.1.3. High Repetitive Accuracy and Automatic Calibration

Generally, to efficiently use a calibrated camera in different field applications where it experiences different environmental conditions such as different temperatures, the camera first needs to be calibrated under those conditions. For example, after calibrating the camera under different temperature conditions, we can analyse the effects on the calibration and consider them in the calibration model. These tests are performed to model possible housing deformation or changes in optical behaviour of a camera in temperatures e.g. from -40 °C up to 85 °C. Therefore, other than simply calibrating a camera under laboratory conditions⁹, a test stand calibration solution needs to be able to calibrate the camera under different continuously-changing environmental and temperature conditions. The calibration should deliver valid and (acceptably) non-correlated calibration parameters which means a significant change in a calibration parameter could be detected and measured.

Based on the above discussion, we can say that a very important requirement of a test stand is the high repetitive accuracy which means the calibration should not be considerably affected by the correlation between the intrinsic parameters and also the extrinsic orientation. As discussed in [2.5.2], the three-dimensional depth information available in bundle-block calibration is very beneficial to decrease the correlation. However, a bundle-block calibration is practically not feasible under the above extreme temperatures and environmental conditions. Therefore, in order to achieve a high calibration accuracy and low parameter correlation in a test stand calibration, we need to simulate the depth-information features of bundle-block calibration in our test stand. This design criterion is further studied in test stand design in [6.2.3].

A fast and automatic calibration (Abraham & Hau, 1997) is required for sustaining certain efficiency and operation capability when dealing with huge number of cameras. Therefore it is required to realize automatic calibration by taking advantage of automation techniques in software and hardware e.g. in image processing algorithms [2.5.3]. Using machines for handling brings an operational automation into our process and further increases the stability and repetitive accuracy of the calibration. This is done by ensuring the accurate positioning of the camera in the ideal position, and for example eliminates thermal interaction with any external objects. This high repetitive accuracy feature is very important especially when we are using the test stand for tests such as alignment testing. It is also important when we want to studying different environmental effects on the camera and performing hardware evaluation such as temperature tests.

⁹ For example: DIN EN ISO 1:2016

6. Hardware Requirements and Test Stand Design

In the previous chapter, we reviewed the general requirements of test stand design. In this chapter - based on those requirements - we introduce new generations of test stands to calibrate the perspective front-view and omnidirectional surround-view cameras benefitting from state-of-the-art designing and manufacturing techniques. These test stands allow us to perform the camera calibration in single-shot calibration in just few seconds, deliver a high calibration accuracy which is comparable to bundle-block calibration accuracy, and provide a very high repetitive accuracy which is necessary for many hardware environmental evaluation applications.

6.1. Specifications of Cameras under Calibration

In order to design test stands that are compatible with one or more special cameras, we require to know their optical specifications. We need to consider the specifications of these cameras when designing our calibration test stands to ensure the full compatibility between the test stand, camera, and the calibration model and to deliver the highest calibration quality and accuracy.



a	b
---	---

 Fig. 6.1. Example of perspective cameras needed to be calibrated
 a: MPC camera, used as a mono front-view camera in driver-assistant systems
 b: SVC camera, used as a stereo front-view camera in driver-assistant systems, using the same optical module as of MPC. (Courtesy of Robert Bosch GmbH)

First we consider the perspective/central-projection cameras. As visible in Fig. 6.1, the perspective cameras used in these calibration efforts are the MPC (Multi-Purpose Camera) and SVC (Stereo-Video Camera). As discussed in [2.3], these cameras are used as front-view cameras in driver-assistant systems and are installed behind the vehicle's front windshield. They feature an integrated processing unit (ECU), a 1280x720 active high dynamic-range CMOS array, and 45° horizontal field-of-view. The specifications of the MPC or SVC optical module (lens and sensor) are provided in Table 6.1.

Camera Parameter	Value
Image Sensor	CMOS
Image Height	720
Image Width	1280
Pixel Size	3.75 [μm]
Colour Filter Array (CFA) Pattern	RCCB (French Pattern)
Nominal Focal Length	5.2 [mm]
Spectral Range (SR)	400-700 [nm]
Depth of Field	≈ 1.5 [m] to Infinity
Lens Diameter	12 [mm]
HFOV	≈ 45 [deg]
VFOV	≈ 24.9 [deg]
Spatial Resolution	≈ 20 [deg/pix]

Table 6.1. Specifications of the test-cameras MPC (Multi-Purpose Camera) used in driver-assistant systems

Next, we review the specifications of the omni-directional cameras which need to be calibrated. As visible in Fig. 6.2, the camera used in the following tests is the NRC (Near-Range Camera) used as a surround-view camera in driver-assistant systems. The specifications of the NRC camera are provided in Table 6.2.



Fig. 6.2. NRC camera, used in surround-view driver-assistant systems
(Courtesy of Robert Bosch GmbH)

Camera Parameter	Value
Image Sensor	CMOS
Image Height	960
Image Width	1280
Pixel Size	3.75 [μm]
Colour Filter Array (CFA) Pattern	RGGB (Bayer Pattern)
Nominal Focal Length	5.2 [mm]
Spectral Range (SR)	400-700 [nm]
Depth of Field	≈ 0.3 [m] to Infinity
HFOV	≈ 192 [deg] (>191 [deg])
VFOV	≈ 144 [deg]
Spatial Resolution	7 [deg/pix]

Table 6.2. Specifications of the test-cameras NRC (Near-Range Camera) used in driver-assistant systems

6.2. Pre-Analysis and Test Stand Design Criteria

We refer to pre-analysis as the operation performed while designing an adjustment network which helps us to make decisions about network geometry and the accuracy of observations to ensure the required accuracy and significance for unknown parameters. Further discussion available in [11] and [12]. We study the design criteria of new test stands incorporating features discussed in [5.1.3], [5.1.1] and [5.1.3]. First, based on the specifications in [6.1], we calculate measurement accuracy parameters in image space. After that, we proceed with the geometrical design and estimation of the necessary depth information for reduction of calibration parameter correlation. After that, using the conceived design parameters, we calculate the measurement accuracy of the test stand in 3D space to fulfil the targeted final calibration accuracy.

6.2.1. Pre-Analysis: Test Stand Point Distribution

As discussed in [5] we need to implement certain features of multiple-view test-field-based calibration as point distribution in our test stand. In order to calculate the point distribution for the bundle-block calibration, we set the image height value (observed point radius) for these points as observations. To calculate distribution factors, we set the radial distance from principal point or correspondingly the sight angle for each consecutive point as new observation and then calculate the PDF based on these values. For this purpose identical radial observations which are image points closer than 1 Pixel were omitted from the dataset which is the threshold for point distribution in image space. By avoiding the recurrent points, we were able to calculate the distribution parameters for bundle-block calibration of our fisheye cameras.

The distribution of the points measured as image radius for bundle-block calibration is very close to a uniform distribution almost up to the edge. This means that a uniform point distribution is present when performing bundle-block calibration. The geometry of multiple-view test-field-based calibration helps us to have points with different image heights and sight angles similar to uniform spacing or uniformly distributed points. Assuming a normal polynomial-fitting (for simplifications) this would mathematically help us to better fit our model to the projection curve (Guest, 1958) (Benhenni & Degras, 2011). Therefore by having points that are uniformly distributed over radius, we will be able to estimate and model the lens distortion effectively.

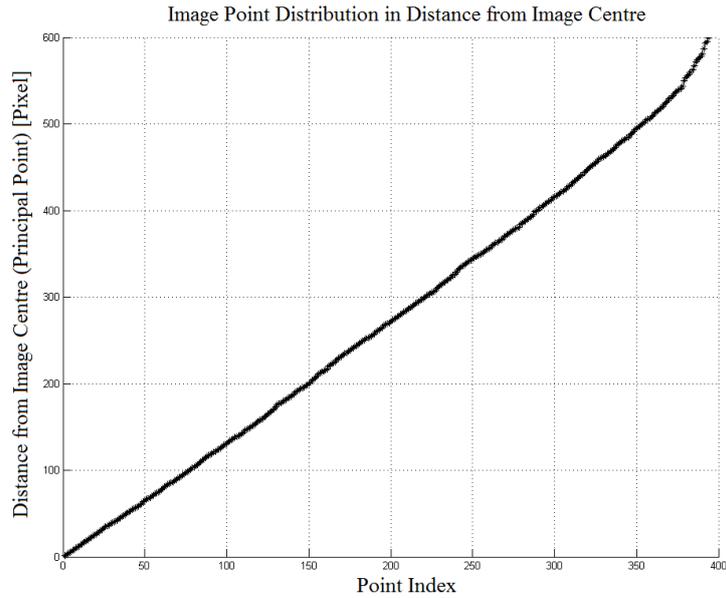


Fig. 6.3. Distribution of point distances in Bundle-Block calibration of fisheye Cameras
(Courtesy of Robert Bosch GmbH)

In Fig. 6.3 we can see the ideal point distribution is almost a straight line and looks very similar to a uniform distribution. Considering a normal distribution for the distance between consecutive points, we get $\mu = 1.6902$ [pixel], $\sigma = 1.7069$ [pixel]. We can see that it has a standard deviation almost equal or not bigger than the expectation for the point distances. The standard deviation shows the quality of point distribution over the whole image and is almost equal to the mean distance of consecutive points. Furthermore, as visible in Fig. 6.3, As a measure of probability distribution of points, by dividing the measurement domain to N parts, the probability of measurements happening in each part is around $\frac{1}{N} \cdot 100\%$. We can conclude that the probability of measurements happening in any image area is almost equal.

6.2.2. Pre-Analysis: 2D Image Measurement Accuracy

As discussed previously, it is essential to calculate the accuracy requirements and their meaning in image space. Here we perform the test stand pre-analysis in order to determine the required accuracy for camera calibration and test stand measurement. As displayed in Fig. 6.4, we can consider a front-view stereo camera with a resolution of 1280x720 pixel and a HFOV of around 45 [deg].

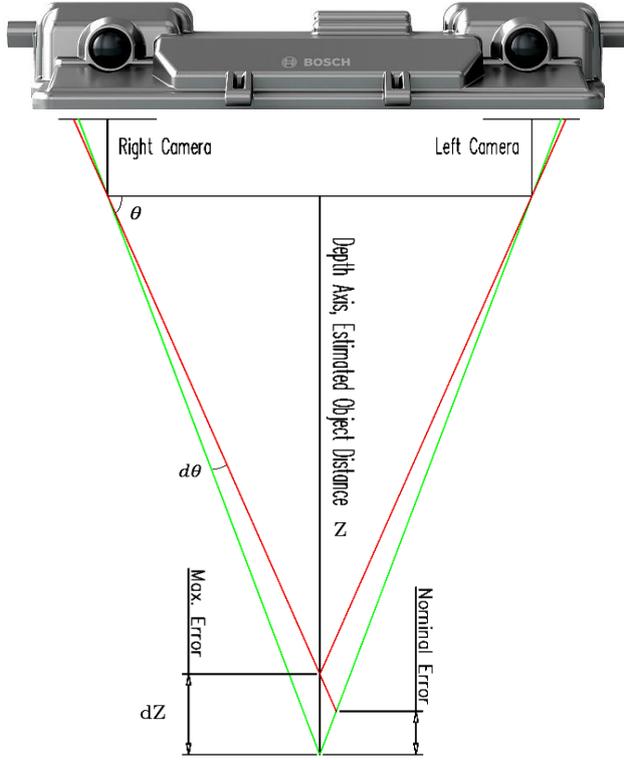


Fig. 6.4. Two-dimensional illustration of error in depth measurement based on measurement accuracy in image.

$$\begin{aligned}
 d\theta &= \operatorname{atan}\left(\frac{Z}{\frac{\text{Base}}{2}}\right) - \operatorname{atan}\left(\frac{Z - dZ}{\frac{\text{Base}}{2}}\right) && \text{Eq. 6.1} \\
 &= \operatorname{atan}\left(\frac{20000}{60}\right) - \operatorname{atan}\left(\frac{20000 - 500}{60}\right) \\
 &\approx 0.0044 \text{ [deg]}
 \end{aligned}$$

In this camera each pixel corresponds to spatial angular resolution of around $\frac{1}{28}$ [deg] and the stereo base distance is around 120 [mm]. We consider Z to be object distance in vehicle forward direction axis, dZ to be the distance error based on stereo image coordinates error, and θ to be the instance angle to the camera base axis. By having a required accuracy of 0.5 [m] in 20 [m] for the distance of reconstructed object, we can approximately calculate the necessary accuracy of the measured coordinates in image space on a two dimensional plane. As displayed in Fig. 6.4, and Eq. 6.1, for a stereo camera with $Z= 20$ [m], $dZ= 0.5$ [m], the respective angular error $d\theta$ corresponding to the depth error will be calculated. This angle equals to 0.1232 [pix]. Again, for a distance of $Z=5$ [m] and a required distance measurement accuracy of around 30 [mm], we calculate a required image coordinate accuracy of around 0.1162 [pix]. Therefore an accuracy threshold of 0.1 [pix] for image coordinates would satisfy the distance or depth estimation accuracy requirement in a real-world scenario. We need to say that here, the value of distance measurement error in the previous calculation only considers the general image coordinates accuracy. Of course, the stereo object reconstruction depends on many other parameters such as the accuracy of relative, extrinsic parameters and so on, but the accuracy of image measurement is of the very first requirements that need to be fulfilled to satisfy the accuracy requirement of stereo modelling and distance measurement. Based on the above calculation, we can say that we need to enforce this accuracy limit - or harder limits - in our camera calibration as one of the main goals to achieve.

6.2.3. Pre-Analysis: Test Stand 3D Depth Configuration

We can calculate, or predict some of the calibration uncertainty measures based on information such as test stand geometrical design parameters and image measurement accuracy. These uncertainty measures consist of, for example, the correlation values between the camera intrinsic and extrinsic parameters. These are the values, which we will have after the camera calibration and try to predict at the pre-analysis stage. The results could be used to improve the test stand design, or to predict the accuracy improvement by design changes. The following analysis, describe how to derive correlation parameters from the relation between camera extrinsic and intrinsic parameters during camera calibration. Here we focus on the correlation between the focal length change ($c2-c1$) and external translation ($z2-z1$). After we derive the formulas, we would use them to compare the correlation parameters for the current/existing test stand and the new test stand design configuration.

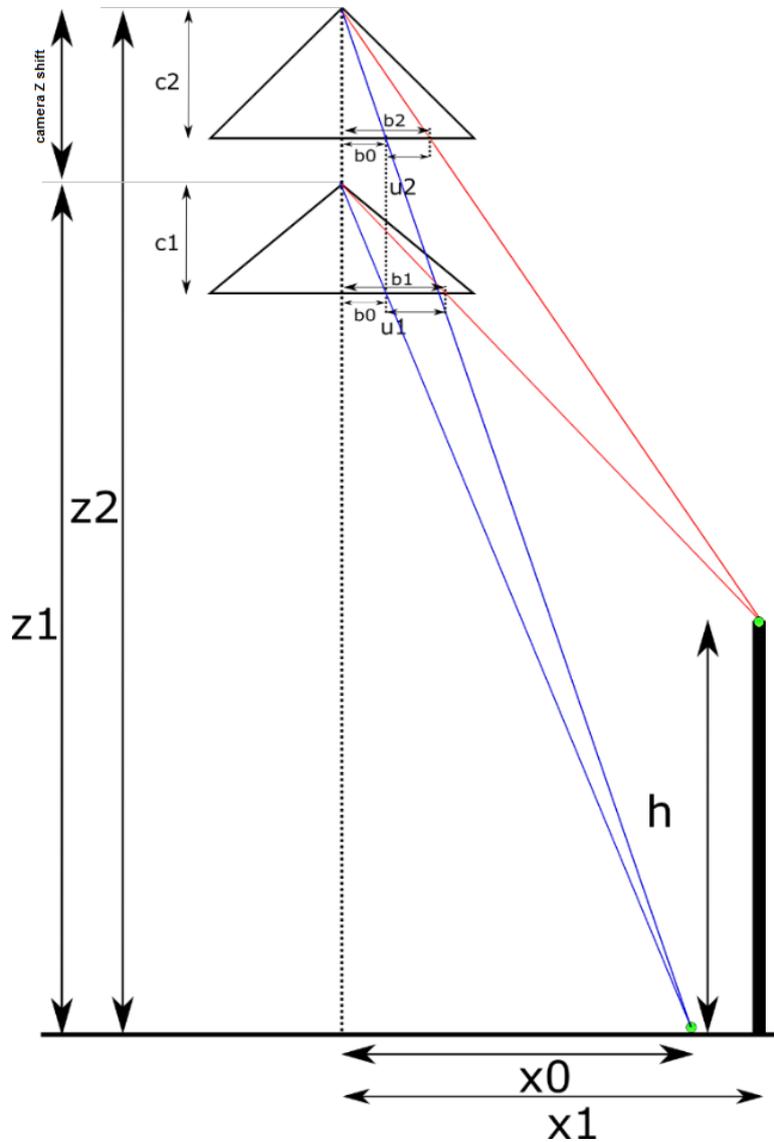


Fig. 6.5. Illustration of coupling between the intrinsic and extrinsic parameters and the effects of stamp points. The shift in camera focal length and the value z are also shown.

We can assume that the measurements in image space are in pixel units and the measurements in object space are in meters. As the parameters are visible in Fig. 6.5, we write the definitions of the correlation parameters based on internal/external camera parameters:

$$z_2 - z_1 = E_{\text{ext}} \quad , \quad c_2 - c_1 = E_{\text{int}} \quad , \quad u_2 - u_1 = E_{\text{bv}} \quad \text{Eq. 6.2}$$

In Eq. 6.2, the parameter E_{ext} is the estimation error in camera extrinsic parameters in the z-axis direction, E_{int} is the estimation error in camera intrinsic parameters in the optical-axis direction, and E_{bv} is the image measurement accuracy. The primary parameter relations are the followings:

$$\frac{b_0}{c_1} = \frac{x_0}{z_1} \quad , \quad \frac{b_1}{c_1} = \frac{x_1}{z_1 - h} \quad , \quad \frac{b_0}{c_2} = \frac{x_0}{z_2} \quad , \quad \frac{b_2}{c_2} = \frac{x_1}{z_2 - h} \quad \text{Eq. 6.3}$$

Now we derive the coupling condition:

$$\begin{aligned} x_0 = \frac{b_0 z_1}{c_1} = \frac{b_0 z_2}{c_2} &\Rightarrow \frac{z_1}{c_1} = \frac{z_2}{c_2} \\ \Rightarrow \boxed{z_1 c_2 - z_2 c_1 = 0} &: \text{Coupling conditions} \end{aligned} \quad \text{Eq. 6.4}$$

And also the coupling proportionality is derived:

$$\begin{aligned} c_1 - c_2 = \frac{b_0 z_1}{x_0} - \frac{b_0 z_2}{x_0} = \frac{b_0}{x_0} (z_1 - z_2) = \frac{c_1}{z_1} (z_1 - z_2) \\ \Rightarrow \boxed{E_{\text{int}} = \frac{c_1}{z_1} E_{\text{ext}}} &: \text{Coupling proportionality} \end{aligned} \quad \text{Eq. 6.5}$$

In Eq. 6.3, Eq. 6.4 and Eq. 6.5 the parameters h , x_0 and x_1 are the 3D depth information. They represent depth differences and positions in object space for an object point. Now we draw the coupling errors from the geometrical characteristics of the test stand. With the above equations, we continue to calculate the relation between the internal and external calibration error and the image processing error E_{int} , E_{ext} and E_{bv} :

$$\begin{aligned} E_{\text{bv}} = u_2 - u_1 = (b_2 - b_0) - (b_1 - b_0) = b_2 - b_1 \\ \Rightarrow E_{\text{bv}} = \frac{c_2 x_1}{z_2 - h} - \frac{c_1 x_1}{z_1 - h} = \frac{c_2 x_1 z_1 - c_2 x_1 h - c_1 x_1 z_2 + c_1 x_1 h}{(z_2 - h)(z_1 - h)} \end{aligned} \quad \text{Eq. 6.6}$$

As we know based on the coupling condition:

$$c_2 z_1 - c_1 z_2 = 0 \quad \Rightarrow \quad E_{\text{bv}} = \frac{x_1 h (c_1 - c_2) - x_1 \overbrace{(c_2 z_1 - c_1 z_2)}^{=0}}{(z_2 - h)(z_1 - h)} \quad \text{Eq. 6.7}$$

Based on Eq. 6.7 we will have:

$$\begin{aligned} E_{\text{bv}} = \frac{x_1 h (c_1 - c_2)}{(z_2 - h)(z_1 - h)} = \frac{x_1 h \frac{c_1}{z_1} (z_1 - z_2)}{(z_2 - h)(z_1 - h)} = \frac{x_1 h c_1 (z_1 - z_2)}{z_1 (z_2 - h)(z_1 - h)} \\ z_2 - z_1 = E_{\text{ext}} \Rightarrow E_{\text{bv}} = \frac{x_1 h c_1 (-E_{\text{ext}})}{z_1 (E_{\text{ext}} + z_1 - h)(z_1 - h)} = \frac{x_1 h c_1 (-E_{\text{ext}})}{z_1 (z_1 - h)^2 + z_1 E_{\text{ext}} (z_1 - h)} \end{aligned} \quad \text{Eq. 6.8}$$

Eq. 6.8 gives the relation between the extrinsic coupling error, image processing error and 3D depth information. We now have the general first-order equation based on the image error:

$$E_{bv} z_1(z_1 - h)^2 + E_{bv} z_1 E_{ext}(z_1 - h) + E_{ext} x_1 h c_1 = 0 \tag{Eq. 6.9}$$

We can now calculate the camera extrinsic error from the image measurement accuracy and the test stand physical parameters:

$$E_{ext} = \frac{E_{bv} z_1(z_1 - h)^2}{-E_{bv} z_1(z_1 - h) - x_1 h c_1} \quad [\text{mm}] \tag{Eq. 6.10}$$

Using Eq. 6.10, we can calculate the error in the intrinsic calibration parameters E_{int} based on E_{bv} and the test stand geometrical configuration parameters:

$$E_{int} = \frac{c_1}{z_1} E_{ext} = \frac{E_{bv} c_1(z_1 - h)^2}{-E_{bv} z_1(z_1 - h) - x_1 h c_1} \quad [\text{Pixel}] \tag{Eq. 6.11}$$

Now we use the Eq. 6.10 and Eq. 6.11 to perform correlation-analysis based on the parameters of a standard test stand and compare the results to those from a new test stand with different design parameter. We proceed to calculate the accuracy and correlation measures for an existing/standard test stand shown in [2.5.3]. We calculate the E_{ext} and E_{int} values using the approximate geometrical and depth information of the test stand are given in Table 6.3.

Design Parameters	z_1	c_1	h	x_1	E_{bv}
Value	1200 [mm]	1440 [Pixel]	300 [mm]	55..80 [mm]	0.2 [Pixel]

Table 6.3. Design Parameters of the existing test stand used for calibrating central-projection cameras. The equivalent focal length is given in pixel units as discussed in [2.4.3.1]

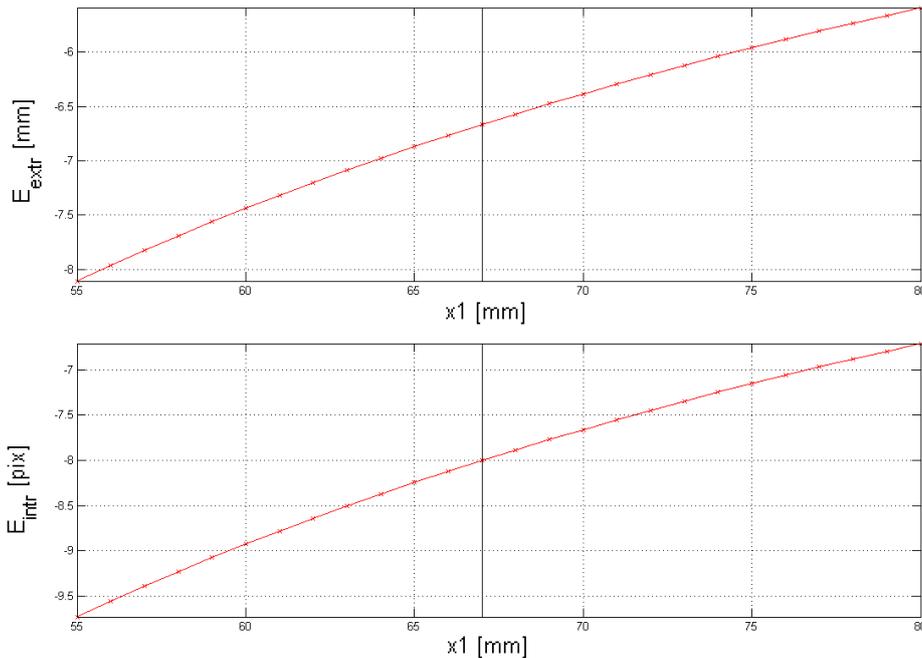


Fig 6.6. Correlation between parameters in the standard test stand

In Fig 6.6, the two curve lines illustrate the possible amount of correlation between the intrinsic and extrinsic parameters based on a varying depth information (x_1 axis) in a standard test stand [2.5.3]. The vertical line shows the actual value of x_1 , and the intersection point with the curve gives the actual correlation value on the vertical axis. Now we consider some of the design parameters of our new test stand which are given in Table 6.4 to calculate the correlation parameters, compare them with the standard test stand and see the improvements.

Design Parameters	z_1	c_1	h	x_1	E_{bv}
Value	1000 [mm]	1440 [Pixel]	240 [mm]	400..500 [mm]	0.2 [Pixel]

Table 6.4. Design Parameters of the new test stand for calibrating central-projection cameras

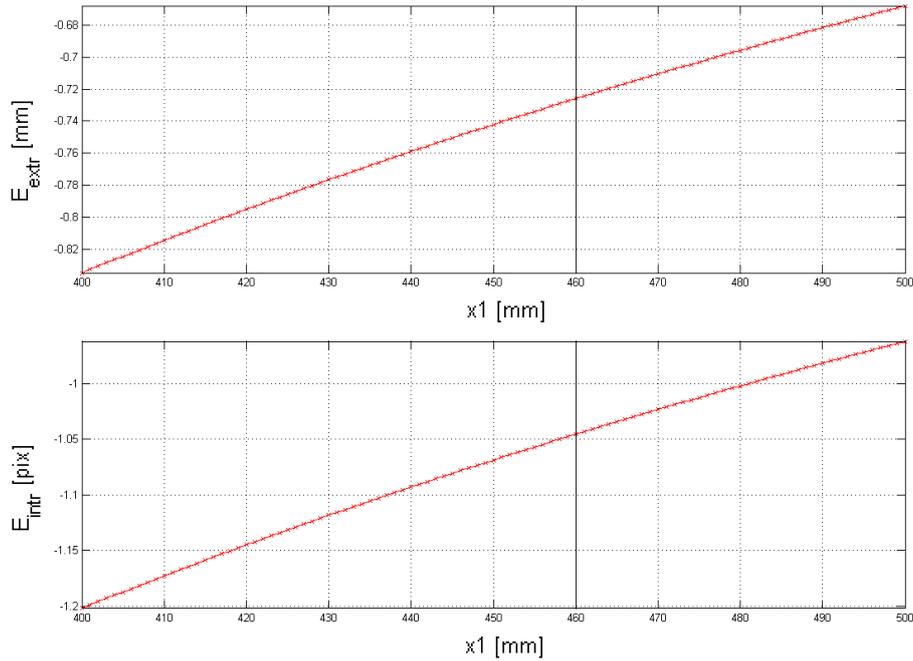


Fig. 6.7. Correlation between parameters in the new test stand

In Fig. 6.7 (as in Fig 6.6), the two curve lines show the amount of correlation between the intrinsic and extrinsic parameters based on a varying depth information (x_1 axis). The vertical line is the actual value of x_1 , and the intersection point gives the actual correlation value on the vertical axis.

Finally, we can compare the results of the pre-analysis from our new test stand with the results of the current standard test stand reviewed in [2.5.3] and illustrated in Fig. 2.26. We can see an improvement of about 925% in accuracy of the extrinsic parameters and about 800% in accuracy of intrinsic parameters (correlation) which are both very significant. It is possible to perform some fine-tuning and further modify the design parameters to get even better results. The design-details of the new test stand is discussed in chapter [6.3.1].

6.2.4. Pre-Analysis: 3D Test Stand Measurement Accuracy

Having calculated the required accuracy in image measurement in [6.2.2], we can continue with the design process. With an approximate design of the test stand i.e. three dimensional depth information from [6.2.3], and the expectation for the calibration accuracy in image space, we can calculate the required accuracy of the observations in 3D space and determine measurement accuracy thresholds. The accuracy of the test stand measurement i.e. the three-dimensional object accuracy should be at least within those limits, or higher, so that the accuracy of estimated parameters meets our expectations.

By projecting the required accuracy in image space into the 3D space using the test stand geometrical design and the specifications of the camera optical mapping function we can calculate the required measurement accuracy in object space similar to GSD calculation. This value is highly dependent on the test stand geometry, optical design and sensor resolution, and is measured for each test stand and camera separately. The test stand is then measured/calibrated using accurate measurement techniques (mechanical and optical) with accuracies better than the required accuracy threshold.

For the new perspective-camera test stand and the stereo camera in Fig. 6.4 the test stand is measured with an expected accuracy of 50 [μm] and accuracy threshold of 100 [μm]. The details of this test stand are discussed in [6.3.1]. Also for the new test-stand for omni-directional camera calibration, the test stand measurement is performed with expected accuracy of 50 [μm]. The details of this test stand are discussed in [6.4.1].

We can finally express that 3D measurement threshold guarantees that the errors in 3D space will not affect the quality and accuracy of camera calibration. Furthermore, this ensures us that the remaining errors after calibration, are due to the optical behaviour of the camera lens, and are not originating from 3D space coordinates error. This helps us further develop our calibration model to model the lens distortion much effectively as discussed in [4].

6.3. Test Stand Specifications for Perspective Camera

One of the main drawbacks of some of the test-fields designed to calibrate the central-projection cameras such as (Abraham, 2004), (Mallon & Whelan, 2004) and [2.5.3] is that most of the control points lay on a single plane in space. This means that although there are three dimensional coordinates assigned to them, they could practically be considered as two dimensional points. This may cause difficulties to fully differentiate between the intrinsic and extrinsic camera parameters during calibration. This is especially noticeable when using complex intrinsic models for camera calibration.

6.3.1. Test Stand Design Details

As discussed in [6.2.1] and illustrated in Fig. 6.9 we have designed the control points in this test stand to lie on multiple planes with each having different distances from camera. As shown in previous chapter, this helps us to have (geometrically) real three dimensional points in space and reduce the correlation between the intrinsic and extrinsic camera parameters.

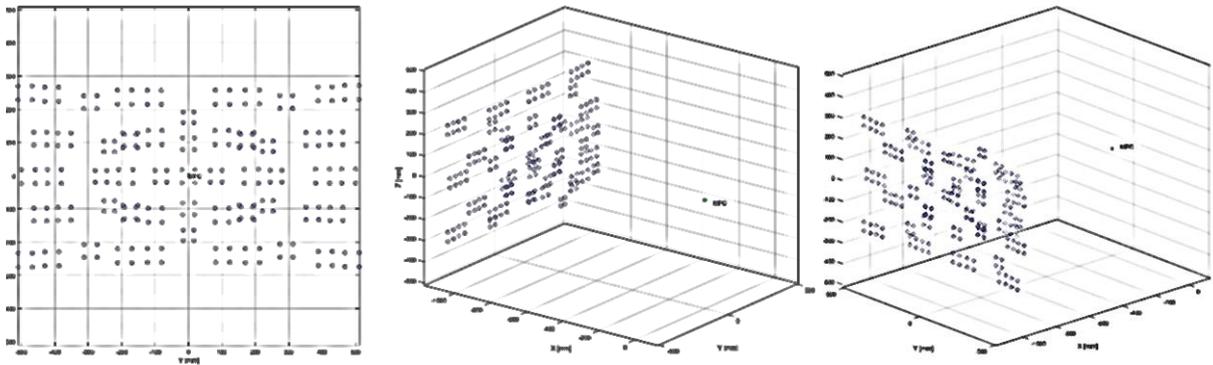


Fig. 6.8. 3D representation of the geomtric design of ME4P2 test stand for central-projection camera calibration
(Courtesy of Robert Bosch GmbH)

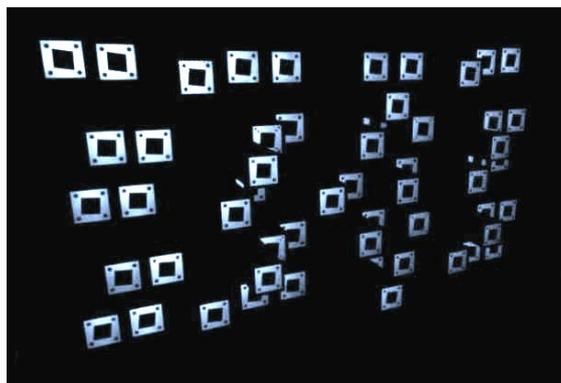
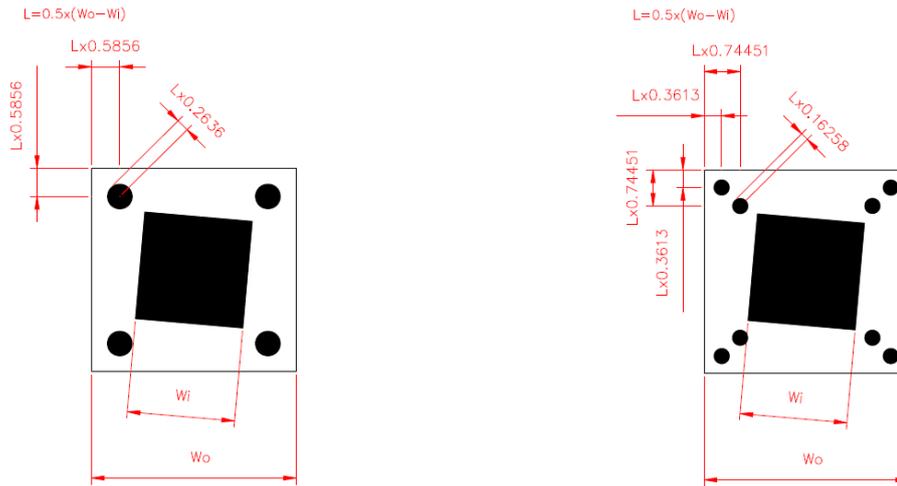


Fig. 6.9. Image of ME4P2 test stand
Control points are visible, which lay on several different planes
(Courtesy of Robert Bosch GmbH)

Furthermore, we have combined the test-fields with slanted edge patterns to be able to simultaneously perform camera sharpness testing i.e. MTF measurement using slanted edge technique (Estribeau & Magnan, 2004). There are also coded point-groups added to the test-field to enable us to perform automatic calibration as also recommended by (Abraham & Hau, 1997).



a	b
---	---

 Fig. 6.10. Representation of the geomtric design of control points in ME4P2 test stand
a: general control points (set of 4 points) b: point groups (set of 8 points)
(Courtesy of Robert Bosch GmbH)

The edges of the points are created by the latest laser-cutting techniques and the coordinates of the control points are measured with laser-based measurement solutions. Complying with the requirement of the pre-analysis in [6.2.4] this measurement technique deliver accuracies around 50 [μm]. Based on the optical specifications in Table 6.1, the GSD – i.e. the sampling distance in object space – in the distance of around 1 [m] is around 800 [μm] and in the distance of 1.5 [m] is around 1300 [μm]. Therefore the measurement accuracy is very smaller than the GSD and is complying with the test stand accuracy requirements discussed in 6.2.2 .

We should mention that a special advantage of this test stand design is the possibility of having operational automation which is the use of machines to perform the calibration procedure and thus avoid any direct human interaction. In comparison with solutions in which human interaction is necessary - which brings unpredicted effects and instability in calibration procedure - using machines instead would increase the stability and accuracy of the calibration and thus ensured the quality of the calibration process.

6.4. Test Stand Specifications for Omni-Directional Camera

The nature of fisheye cameras used in surround-view systems is to have extreme opening angles bigger than 180 degree which means having high spherical projection factor or high radial distortion depending on the interpretation of projection model such as (Mei & Rives, 2007). This implies that the radial distortion is significantly big in comparison with other lens distortion characteristics. Therefore we need to be able to accurately model this distortion. As discussed in [5.1.1], in test-field-based bundle-block calibration we have generally many points that are - because of its multiple view geometry - well distributed over the image and provide us with high calibration quality, and make an accurate distortion modelling possible. Therefore, we need to simulate that point-distribution in the test stand design by understanding the specifications in point distribution of bundle-block calibration solutions. (Nekouei Shahraki & Haala, 2015 B)

6.4.1. Test Stand Design Details

As the first step for the geometrical design of the test stand, we need to study the specifications of the video camera, which is going to be calibrated using this device. The video camera is the NRC (Near-Range Camera) used as a surround-view video system in driver-assistant systems. The specifications of the camera are provided in Table 6.2.

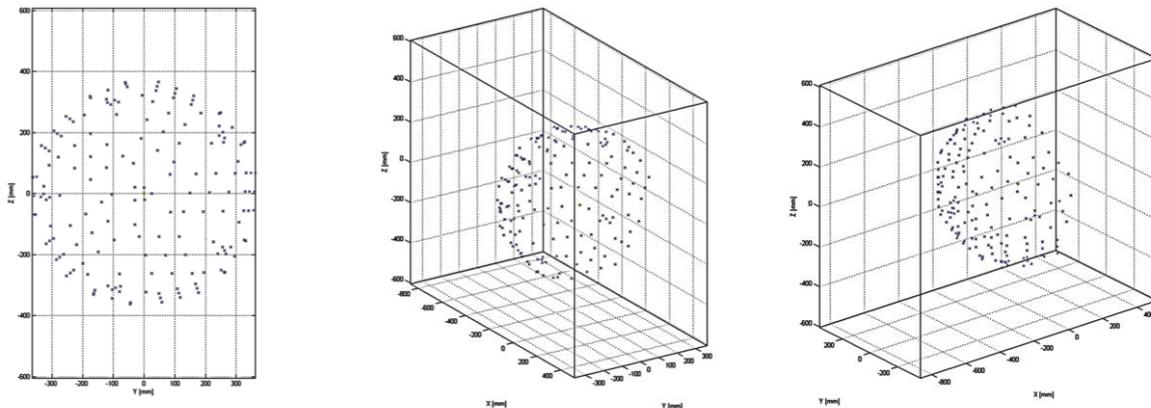


Fig. 6.11. 3D representation of the geomtric design of ME4P3 test stand for fisheye camera calibration (Courtesy of Robert Bosch GmbH)

As discussed in [6.2.1], the ideal bundle-block calibration has a point distribution characteristic similar to a uniform point distribution that increases the camera calibration quality and validity all over the image. To have these advantageous features from bundle-block calibration in our calibration test stand, we have used the geometrical shape of a spherical helix (type: 3D spherical spiral). One of the special characteristics of this spherical helix geometry is that it has a uniform radius which corresponds to the 3D uniform motion of an arbitrary object around an axis (Pottmann, et al., 2002) (Pottmann, et al., 2004) (Nekouei Shahraki & Haala, 2015 B). The positions of the test stand target points are then based on the spherical helix geometry.

In addition, in order to realize an automated calibration process in software - as discussed in [5.1.3] - there are certain coded point-groups added to the test stand control points allowing us to use automated algorithms in the calibration process and thus avoid any direct human interaction. This process is performed by automatically detecting these distinct point-groups and performing point-matching.

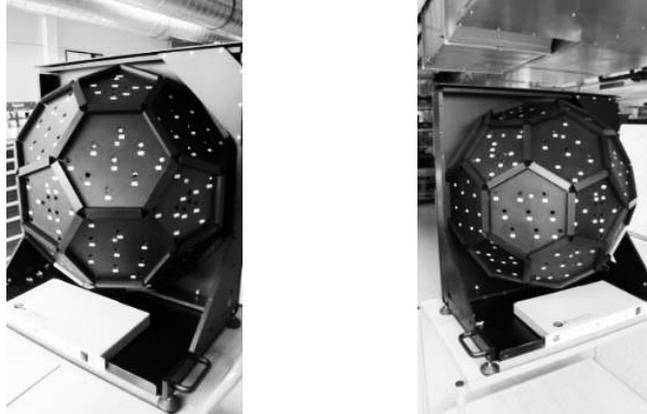


Fig. 6.12. Images of the designed ME4E3 test stand constructed truncated icosahedrons and the designed points on the surface are visible (Courtesy of Robert Bosch GmbH)

The geometrical form was realized using the shape of a sliced 3D truncated icosahedron, which is similar to a regular dodecahedron with hexagonal and pentagonal shapes (Stakhov, 2009) (Hosoya, 2011). This shape practically allows the implementation of a spherical helix on multiple planar surfaces. To keep the ideal designed point distribution in image, the image points are analytically projected on the surfaces of the icosahedron and the initial 3D coordinates are calculated. There are a total of 215 points in test stand, with at least 150 visible points distributed in image as illustrated in Fig. 6.15. The number of visible points depend on the camera sensor active ROI corresponding to the image height and the VFOV.

The design characteristics of the control points were finally achieved by analysing the geometrical features of the test stand and the stray-light testing as the second necessary optical test - as discussed in [5.1.2] - required to be performed on the camera. The points are designed to be active-illuminated diffuser points with an accurate no-shadow design. The centre of diffusor points are measured indirectly using the mechanical touch/arm measurement techniques.



Fig. 6.13. Representation of the geomtric design of control points (target points)
Active illumination with no-shadow design (Courtesy of Robert Bosch GmbH)

All the control point mechanical parts and elements are constructed using precise machinery with accuracy of around 20-50 μm . The test stand is measured mechanically with the approximate accuracy of 50 μm which complies with the requirement of the pre-analysis in [6.2.4]. Again, based on the optical specifications in Table 6.2, the GSD in the distance of 0.5 [m] is around 1250 μm and in the distance of 0.37 [m] is around 925 μm . Therefore the measurement accuracy is very smaller than the GSD and complies with the accuracy requirements discussed in 6.2.2.

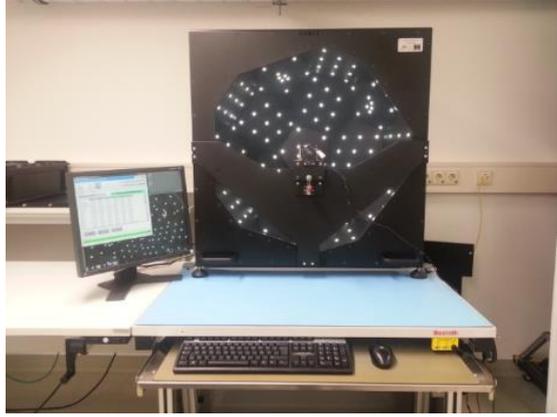


Fig. 6.14. Image of the final functioning ME4E3 test stand in the laboratory
(Courtesy of Robert Bosch GmbH)

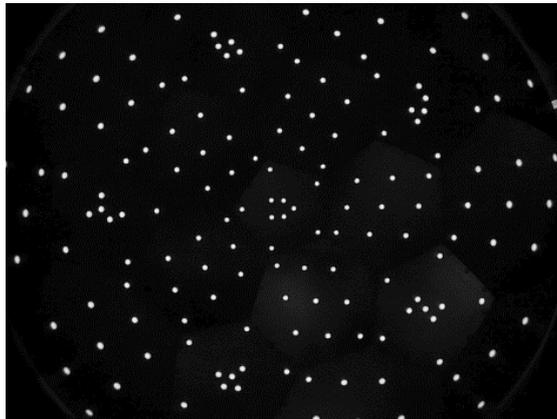


Fig. 6.15. Captured image from ME4E3 test stand with a NRC Camera
(Courtesy of Robert Bosch GmbH)

The final design and the test stand is represented in Fig. 6.14 and Fig. 6.15, showing the achieved geometrical characteristics of projected points in the camera. The projection of the modified three dimensional spherical helix (spherical spiral) is clearly visible in the image.

As discussed in [5.1.1] we should guarantee a certain point distribution to avoid Runge's phenomenon from happening. By studying the 3D geometry of test stand points and image points, it would be clear that the Runge's Phenomenon does not happen. This is because an almost equidistance point distribution similar to equispaced interpolation points is present, a high point density i.e. the number of points far exceeding the degree of the distortion model is achieved. It is also because we are avoiding high degrees of the distortion model (Dahlquist & Björck, 2003) with the current radial distortion polynomial up to the sixth radial exponent. Furthermore, this uniform and dense point distribution would help us to have extrapolation capabilities (Laderman & Laderman, 1982) in the distortion model such as C. Mei further beyond the last detected image point near the image edge.

We can also perform camera alignment testing using this test stand as discussed in [2.5.3]. Thus we are able to evaluate the geometrical alignment of the camera housing and the interior elements such as the alignment of sensor surface compared to the defined optical axis and the reference coordinates system. This is realized by designing an accurate and stable (machine-operated) camera fixture defined as origin of the unified test stand coordinates system. The latter is achieved by using reference points on both test stand and camera fixture.

6.4.2. Design Statistical Analysis and Point Distribution

Having the test stand point coordinates on the surface icosahedron, we can calculate the projected coordinates in camera and analyse the point distribution in image space. We analyse the distance of consecutive points from each other and from the image centre (principal point) and generate the point distribution graph comparable to Fig. 6.3. We can then study the histogram of the point distribution over the whole image to get further information about the point distribution in image space.

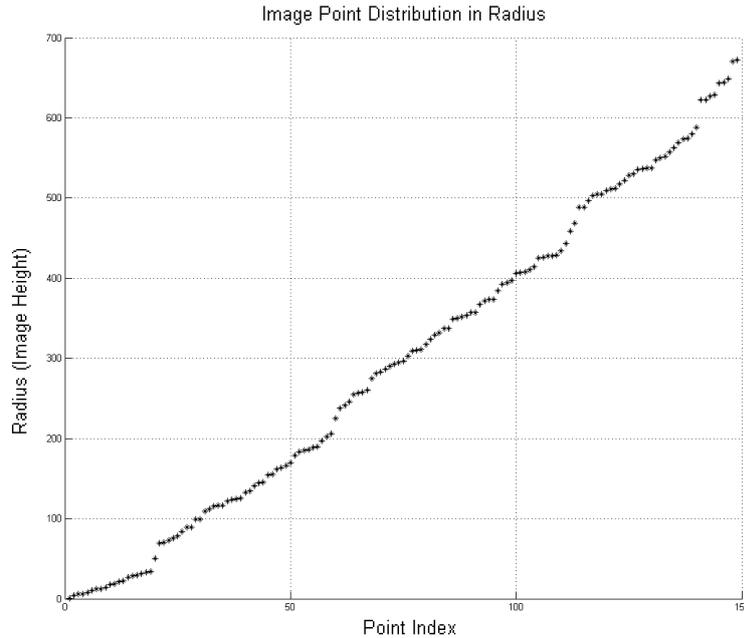


Fig. 6.16. Image point distribution in radius for Test Stand calibration of fisheye Cameras
(Courtesy of Robert Bosch GmbH)

Fig. 6.16 illustrates the test stand point distribution in image by considering each single point (point index) and its corresponding image height. We can see that it is very much similar to a uniform distribution. The small deviations from the straight line are because of some practical and mechanical limits of construction techniques and the introduction of geometrically-coded point-groups for performing automatic point matching in the calibration process.

After analysing the test stand point distribution, we get the values $\mu = 4.5 - 5.4$ [pixel], $\sigma \approx 5$ [pixel] for a normal distribution of the distance difference between consecutive points. This shows compliance with the design criterion discussed in [5.1.1] as the standard deviation is almost equal to the expectation for the point distances i.e. mean distance of consecutive projected points in image.

As we see in Fig. 6.16, and to calculate a measure of probability distribution of points, we can divide the measurement domain to N parts. The probability of measurements happening in each part is around $\frac{1}{N} \cdot 100\%$. This complies with the criteria discussed in 6.2.1 that the probability of measurements happening in any image part is almost equal. We can finally say that this point geometry potentially helps in estimating a valid projection function and modelling the lens distortion accurately all over the image comparable to bundle-block calibration accuracy.

In the next chapters, we use the new designed test stands and perform camera calibration and evaluate the repetitive accuracy of the estimated calibration model.

6.5. Test Stand calibration of Perspective Camera

In the following tests we performed the test stand and the bundle-block calibration which is referred to the multiple-view test-field-based calibration for cameras to be able to compare their repetitive accuracy. In our tests, the bundle-block calibration could not be performed using the designed test stand from [6.3.1] as a test-field. This is because the complex point geometry, as they shadow and block themselves when the camera changes its orientation and position. Therefore we used the same test-field from [2.5.3] to perform the bundle-block calibration. Using the same camera we performed the bundle-block calibration seven times independently and compared the deviations of the estimated calibration models to calculate the final repetitive accuracy.

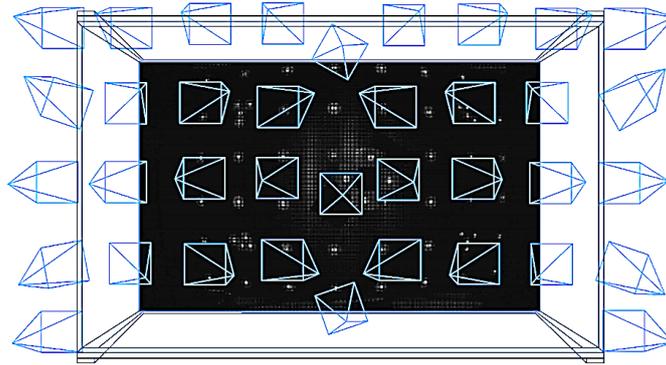


Fig. 6.17. Camera approximate positions in the bundle-block calibration

Each square represents one camera position

In each position, multiple images are taken with different orientations

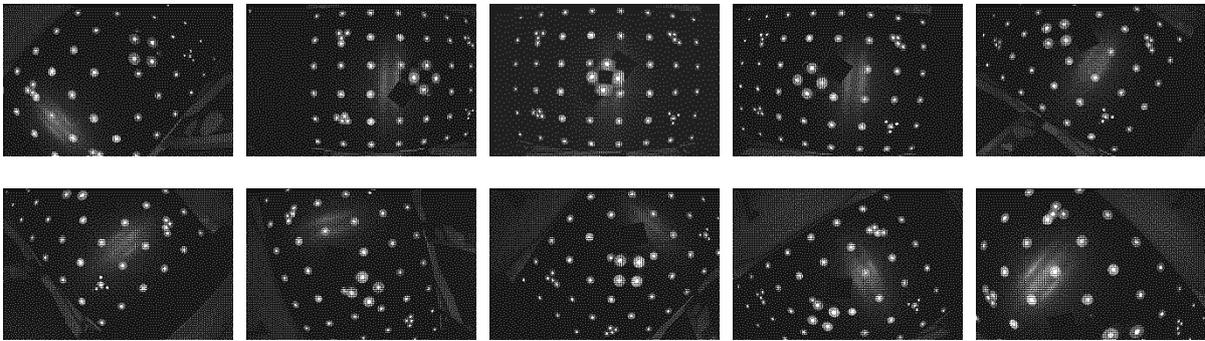


Fig. 6.18. Sample images from bundle-block calibration of MPC camera

Images are taken by continuous changing of orientations and positions

In Fig. 6.20 and Fig. 6.21, the bundle-block configuration is shown. In each bundle-block calibration, around 40 images are used for each camera. This provides us around 1500 observation per calibration from test-field control points (Nekouei Shahraki & Haala, 2015 B). The test stand single-shot calibration of the perspective camera is performed on the designed test stand discussed in [6.3.1] and illustrated in Fig. 6.20.

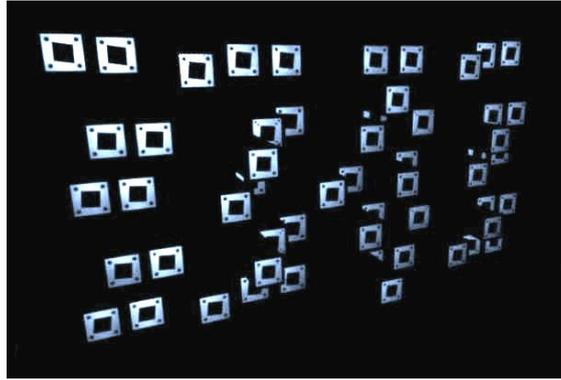


Fig. 6.19. Image of ME4P2 test stand
Control points are visible, which lay on several different planes
(Courtesy of Robert Bosch GmbH)

As discussed in [5.1.3] and [6], one of the most important factors for our calibration solution is the repetitive accuracy of the calibration parameters. In the followings, the accuracy comparisons, is performed by analysing the repetitive accuracy of bundle-block calibration using the planar test-field from Fig. 6.17 and the test stand calibration from Fig. 6.20. When performing each calibration for both of the solutions, the camera has been mounted/placed again in the camera fixture. The calibration model used in this analysis is the Tschebycheff model discussed in [2.4.3.6].

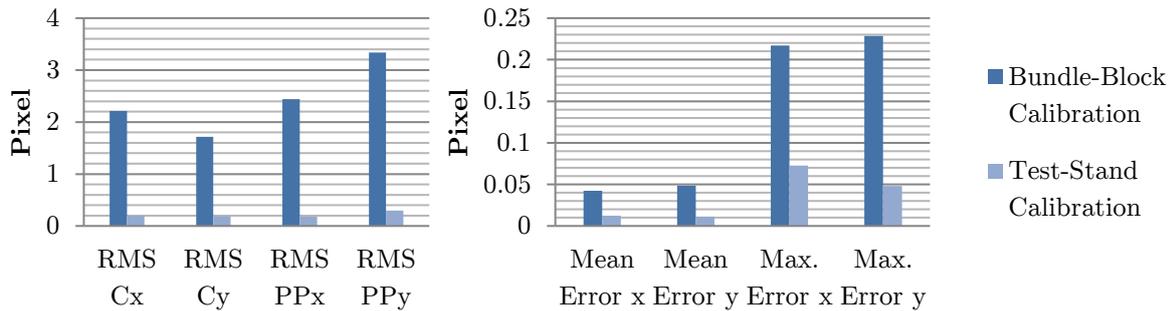


Fig. 6.20. a: Analysis of repetitive accuracy of camera calibration parameters (test stand and bundle-block)
b: Analysis of repetitive accuracy using test stand and bundle-block calibration using a single MPC perspective camera. (Courtesy of Robert Bosch GmbH)

The accuracy of each calibration iteration (RMS) is better than 0.2 pixel for both calibration solutions with maximum error of less than 1 pixel. Fig. 6.20 represents the repetitive accuracy factors taken from processing of each estimated calibration model generated using test stand and bundle-block calibration. The “Mean Error” and “Max Error” values are calculated in image x and y directions respectively from generated distortion-field of each calibration. The latter is estimated by calculating their deviation from each other on defined projected 3D points i.e. a simulated dense object points in space.

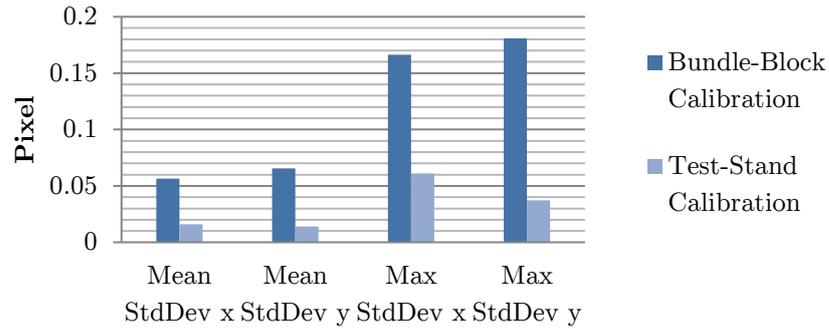


Fig. 6.21. Analysis of the repetitive accuracy of calibrations dataset.
Multiple test stand and bundle-block calibration of a single camera
(Courtesy of Robert Bosch GmbH)

Fig. 6.20 represents the accuracy factors corresponding to the repetitive accuracy taken from processing of the calibration datasets generated using test stand and bundle-block calibration. The “Mean Error” and “Max Error” values are calculated in image x and y directions respectively from generated distortion-field of each calibration. The latter is estimated by calculating their deviation from each other on defined projected 3D points i.e. a simulated dense sphere in space. Fig. 6.21 illustrates the standard deviation of re-projection errors for all of the simulated 3D points as further accuracy measures calculated using the calibration dataset for both of the calibration solutions. A discussion on these repetitive accuracies and their interpretation is available in [06.7].

6.6. Test Stand calibration of Omnidirectional Camera

In our tests we performed the bundle-block calibration which – similar to the one performed previously - is referred to the multiple-view test-field-based calibration for fisheye cameras. This procedure is very similar to the one in previous chapter, and again because of the geometrical characteristic of the test stand, we were not able to use the test stand as the test-field in bundle-block calibration. Therefore we used the test-field discussed in [2.5.2] and Fig. 2.23. We used around 40 images for each camera which is a total of around 1500 observation per calibration (Nekouei Shahraki & Haala, 2015 B).

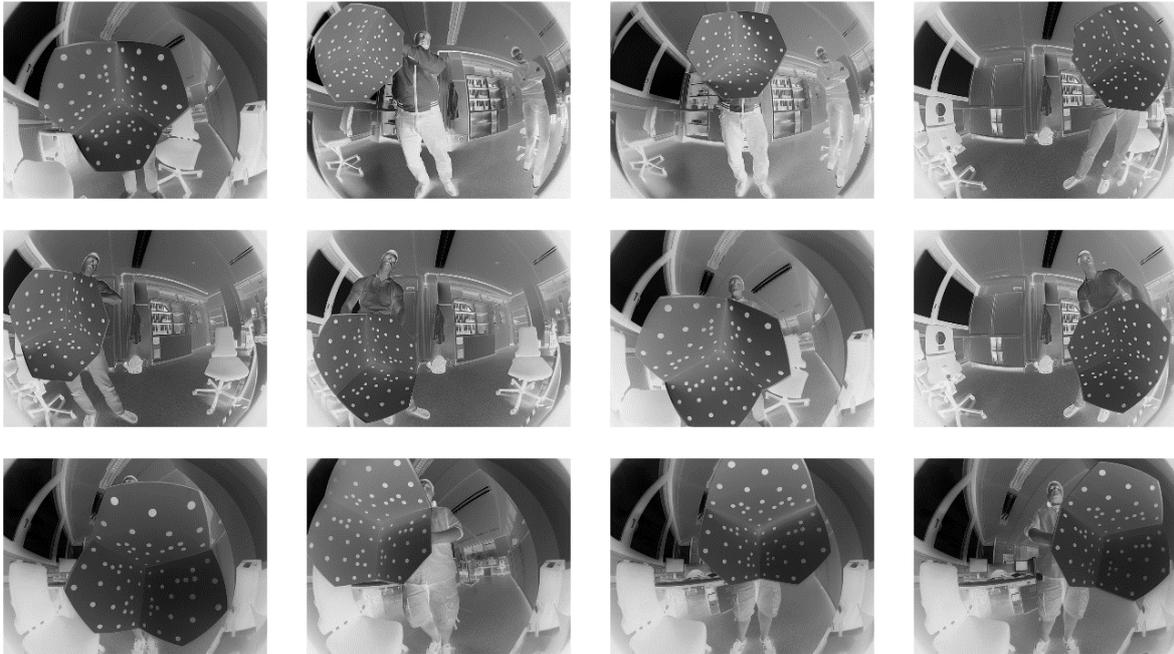


Fig. 6.22. Sample images from bundle-block calibration of NRC camera

Images are taken by continuous changing of orientations and positions of test-field relative to the camera

As also shown in Fig. 6.22, the configuration of bundle-block calibration is similar to the configuration discussed in Fig. 2.24 with the possible camera positions and orientations in between. Using the same fisheye camera we performed the bundle-block calibration seven times independently and compared the results and the deviations of the estimated models to calculate the final repetitive accuracy. The Test-stand calibration of the fisheye camera is performed on the designed test stand discussed in [6.4.1] and illustrated in Fig. 6.23.

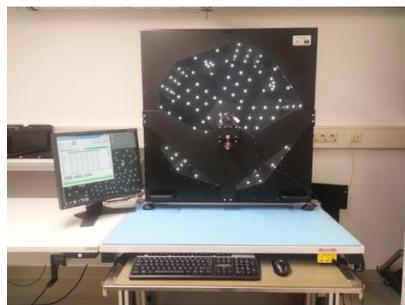
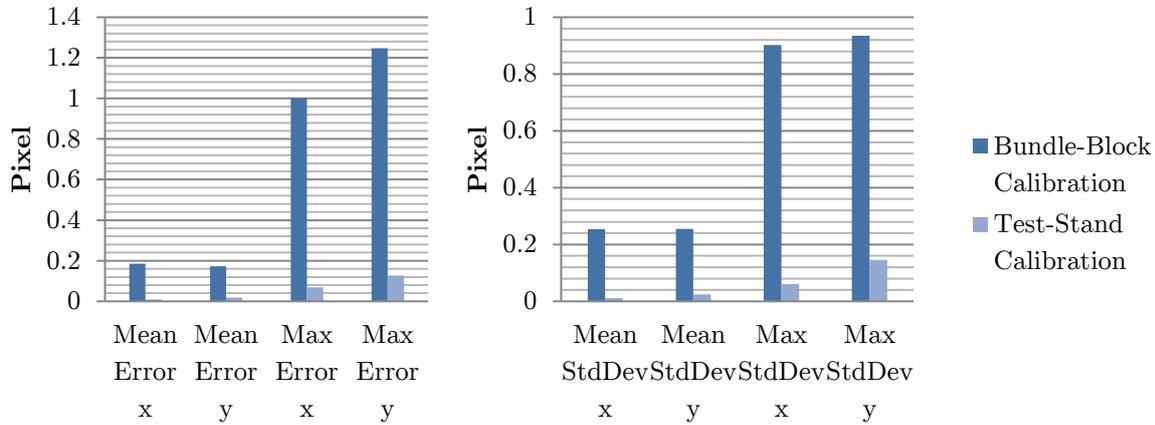


Fig. 6.23. Image of the ME4E3 test stand
(Courtesy of Robert Bosch GmbH)

The accuracy comparisons, is performed by analysing the repetitive accuracy of bundle-block and test stand calibration solutions. Note that between each calibration for both of the solutions, the camera has been mounted/placed again in the camera fixture to include positioning tolerances in our calculation. The calibration model used in this analysis is the extended omnidirectional model discussed in [3.7.1] [3.7.2] [3.7.3].



a	b
---	---

 Fig. 6.24. a: Analysis of repetitive accuracy of test stand and bundle-block calibration for a fisheye NRC camera
b: Analysis of repetitive accuracy of multiple calibrations of test stand and bundle-block calibration for a fisheye NRC camera. (Courtesy of Robert Bosch GmbH)

The accuracy of each calibration iteration (RMS) is better than 0.2 pixel for both calibration solutions with maximum error of less than 1 pixel. In Fig. 6.24a we see the difference in the repetitive accuracy of test stand and bundle-block calibration for fisheye camera calibration. Using the same process described in [6.5], Fig. 6.24 shows the standard deviation of re-projection errors using simulated 3D points, and is calculated using each individual calibration dataset for both of the calibration solutions. A discussion on these repetitive accuracies and their interpretation is available in chapter [6.7].

6.7. Discussion on the Values of Repetitive Accuracy

As illustrated in Fig. 6.21, and discussed in [6.2.3] there is almost always an amount of correlation between the intrinsic parameters such as C_x , C_y and the extrinsic parameters such as translations and angles. These correlations mostly depend on the spatial 3D distribution of points in space (illustrated in Fig. 6.5) and therefore in the bundle-block calibration depend on the block configuration and relative position and orientation of camera and test-field. When calibrating a camera several times, it is usually not always possible to guaranty the realization of the same ideal point distribution in every bundle-block calibration iteration, and thus the deviations or changes in point distributions – to some extents- affect the correlations and the estimated camera intrinsic parameters. In contrast to that, in test stand calibration the presence of a stable 3D point distribution - which could be guaranteed to be the same in each calibration iteration - has contributed to the repetitive accuracy of the test stand calibration. This means that in a repetitive calibration process on the designed test stand, we almost always get the same results and the estimated values of the calibration parameters.

Another reason behind the accuracy difference is that the advantage of an automated procedure in test stand (i.e. performed with machine/robots). This has contributed to a more stable camera position and orientation, which helped maintaining the same 3D point distribution for each calibration iteration, and has further increased the repetitive accuracy and stability.

As also mentioned in [6.5], as a post-calibration analysis, we have calculated the deviation of each calibration set i.e. the estimated calibration model from each other on defined projected 3D points which represent a simulated dense sphere in space. Fig. 6.22 and Fig. 6.24 illustrate standard deviations of re-projection errors for all of the simulated 3D points, and are calculated using each individual calibration result for both of the calibration solutions. We should note that these values are also potentially influenced by the amount of noise or random error in two different calibration iteration. This is mostly caused by a change in point detection accuracy in oblique images with high relative orientations between camera and the test-field in this bundle-block calibration. This again describes the difference in the repetitive accuracy, which is one of the many important aspects in the design and development of these new test stands.

7. Performing Test Stand Calibration using Free-Function Model

In this chapter, we employ the free-function calibration model introduced in [4] to model the camera lens distortion. As said previously as an initial calibration based on the process described in [4.4], [4.4.1] and [4.4.4]. We first calibrate the camera with the designed test stand using a standard distortion model. Then using the residuals calculated on the observed control points we generate the local lens distortion field using the method described in [4.3] and [4.4.2]. After that, the free-function model is used to model the overall lens distortion that contains the local distortion. Finally, after performing the parameter estimation process, the results are used to regenerate the distortion model and compare it with the input data and evaluate the quality of lens distortion modelling.

7.1. Free-Function Calibration on Perspective Camera

Now using a perspective camera from [6.1] (MPC camera) and the test stand calibration method from [6.5], we calibrate the camera initially using the Tschebycheff model [2.4.3.6]. Then the primary and local lens distortion-field using Gaussian interpolation [4.3] are generated.

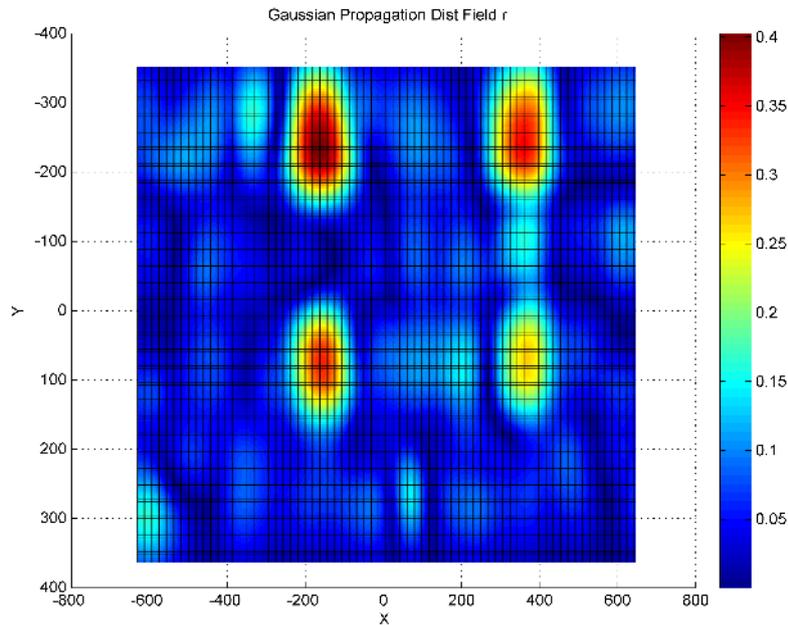
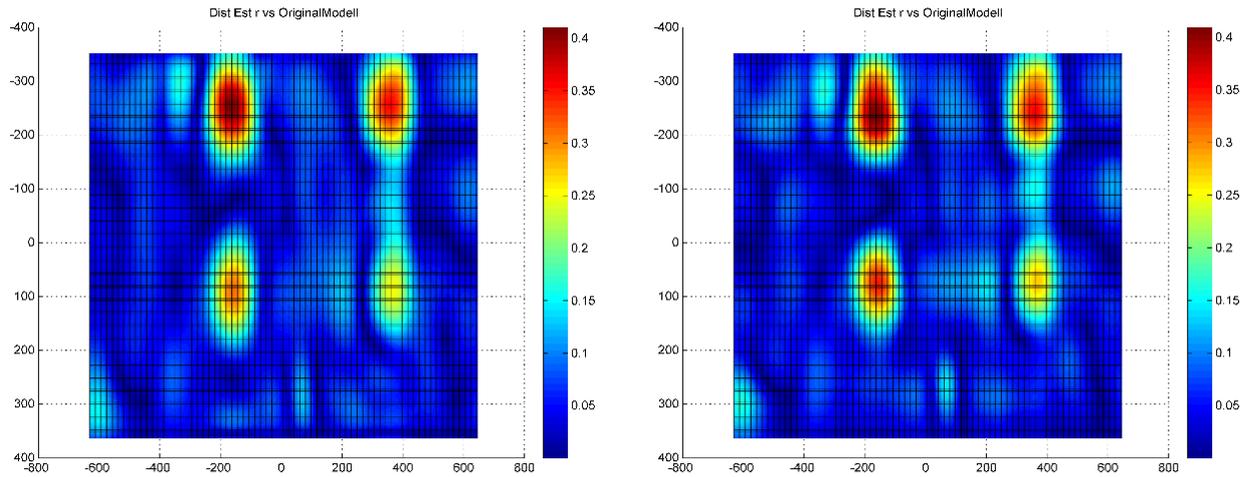


Fig. 7.1. Example of local distortion field (input data) generated using Gaussian-interpolation for a perspective camera. Distribution domain size= 300, StdDev factor= $\frac{1}{5}$. The colour-bar and axes coordinates are in pixel units.

(Courtesy of Robert Bosch GmbH)

Fig. 7.1 shows the amount of the local lens distortion in radial direction. Having these data as input and adding them to the distortion-field generated by the original calibration model, we can perform the free-function model estimation.



a	b
---	---

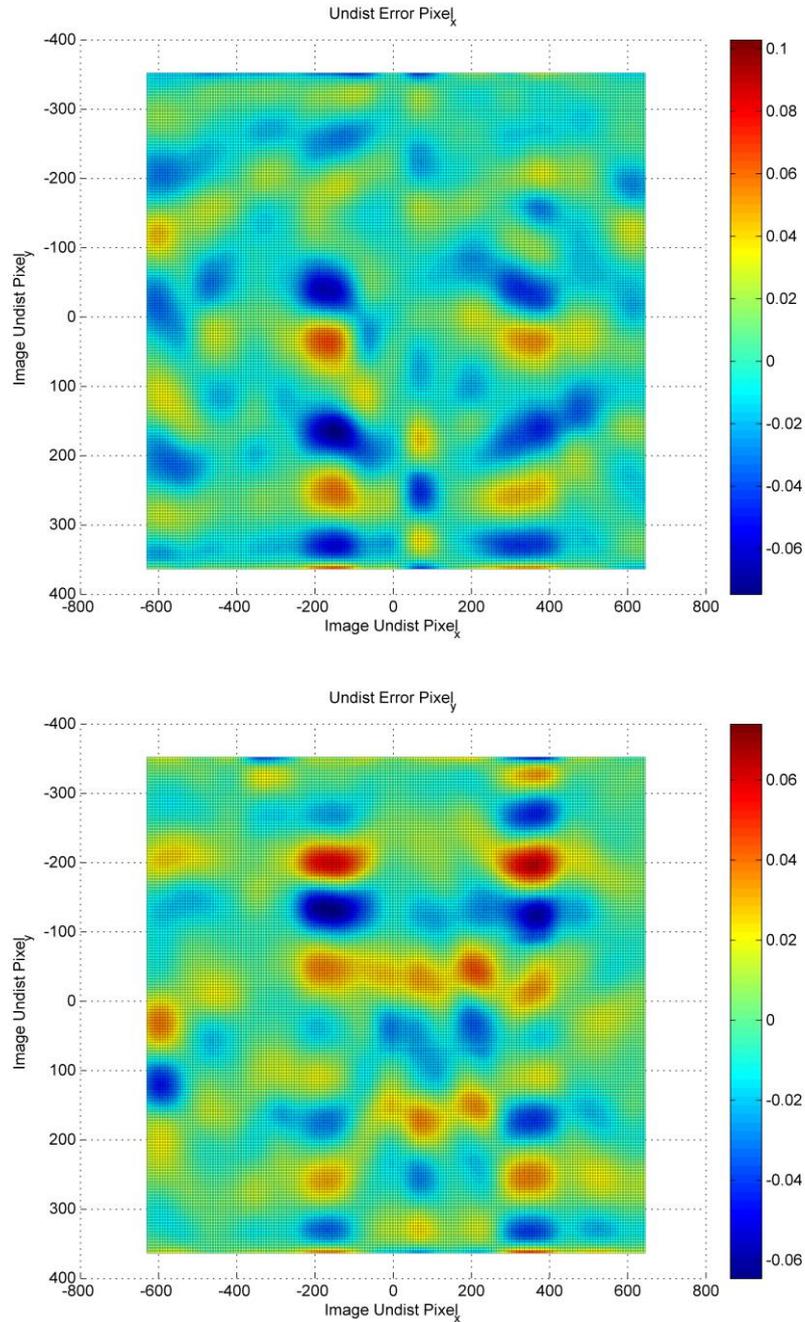
 Fig. 7.2. Result of free-function modelling (modelled local lens distortion) for a central-projection image (in radial direction), Distribution domain size= 300, StdDev factor= $\frac{1}{5}$
a: free-function degree= 8 b: free-function degree= 16
The colour-bar and axes coordinates are in pixel units
(Courtesy of Robert Bosch GmbH)

Fig. 7.2 shows the results of local lens distortion modelling using the free-function model. We conducted distortion modelling using the free-function model of degree 8 and 16 ($n=8$ and $n=16$) and using the distortion-field in Fig. 7.1 and the original model as input data. After that, the calibration is performed following the recommended process of autonomous calibration by S. Abraham (Abraham & Hau, 1997) (Abraham, 2004).

As we have seen above and by comparing the results to Fig. 7.1, the quality of local lens is fairly modelled, and the modelling gets better by using higher degrees for the free-function model. For example, the free-function model of degree 16 has a higher accuracy than the free-function model of degree 8. The accuracy measures are available in the next chapter.

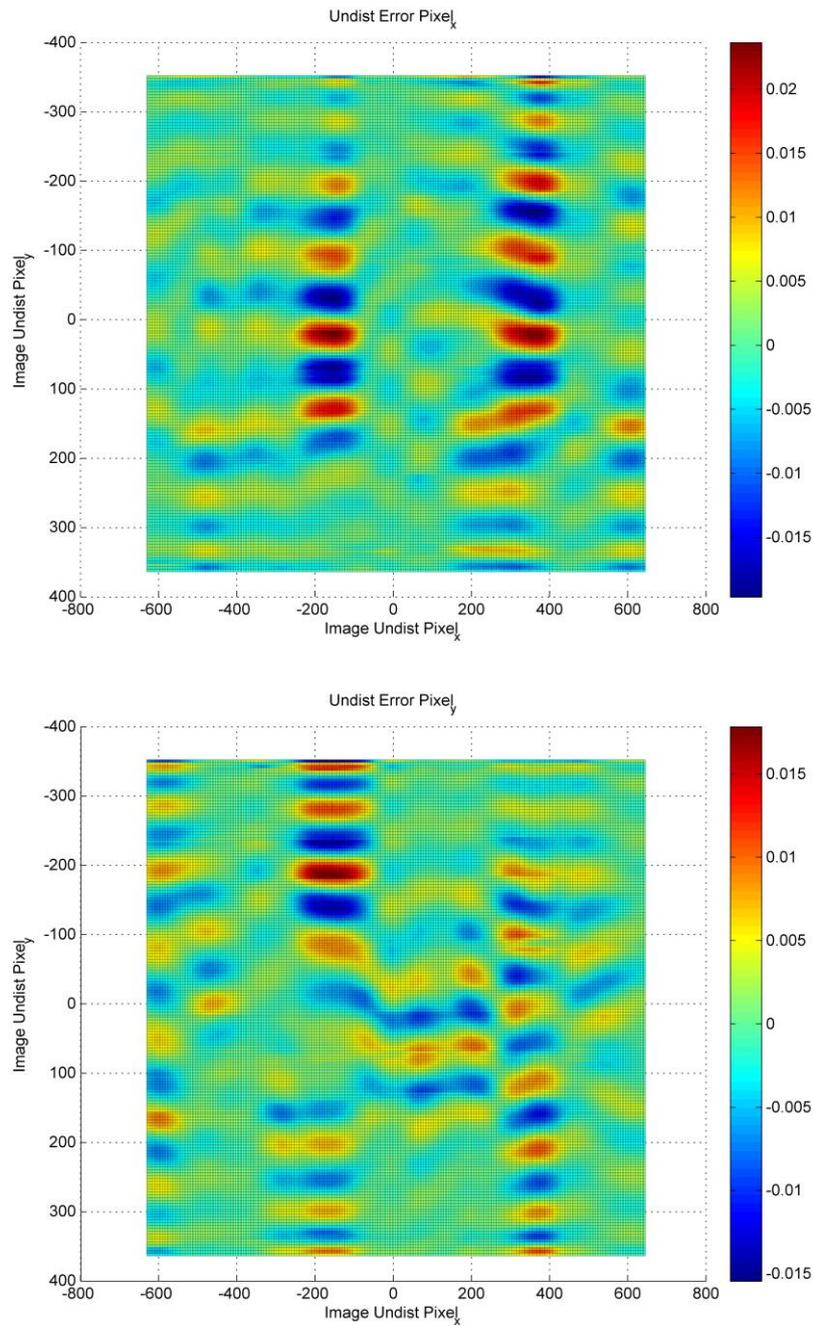
7.1.1. Accuracy and Error Analysis

In this chapter we study the accuracy parameters of the free-function model performed with the degree of 8 and 16 ($n=8$ and $n=16$) and compare them to standard calibration models. Here we analyse the accuracy of free-function using the input data from Fig. 7.1 for a perspective camera. After the distortion is modelled, we calculate the residuals or the difference between the input distortion field and the modelled one to estimate how good the free-function model has fit itself to the distortion field.



a
b

 Fig. 7.3. The remaining errors from free-function calibration for a perspective image Using free-function of degree 8, a: errors in image x direction, b: errors in image y direction,



a Fig. 7.4. The remaining errors from free-function calibration for a perspective image
 b Using free-function of degree 16, a: errors in image x direction, b: errors in image y direction,

Fig. 7.3 and Fig. 7.4 show the remaining errors of local lens distortion modelling over the whole image in x and y directions using the free-function model of degree 8 and 16. We can see that the errors are very small which means the local distortion is very well modelled.

Calibration Model	RMSE x	RMSE y	Max Error x	Max Error y
Brown-Fryer (9 parameter)	0.159	0.149	0.544	0.499
Grün (44 parameter)	0.173	0.159	0.719	0.530
Tschebycheff 3 rd order (18 parameter)	0.173	0.153	0.647	0.512
Legendre (34 parameter)	0.166	0.150	0.723	0.559
FF-Model (degree 4)	0.042	0.029	0.204	0.113
FF-Model (degree 6)	0.023	0.022	0.104	0.089
FF-Model (degree 8)	0.017	0.016	0.080	0.068
FF-Model (degree 10)	0.011	0.013	0.053	0.052
FF-Model (degree 12)	0.008	0.010	0.037	0.044
FF-Model (degree 14)	0.005	0.007	0.024	0.029
FF-Model (degree 16)	0.003	0.005	0.014	0.021

Table 7.1. Accuracy analysis of the free-function model in comparison with standard distortion models
(All the parameters are in pixel units)

Degree of FF-Model	Required Memory (KB)	CPU operations per Pixel
Degree 4	10.2	8
Degree 6	15.4	12
Degree 8	20.5	16
Degree 10	25.6	20
Degree 12	30.7	24
Degree 14	35.8	28
Degree 16	41	32

Table 7.2. Memory requirement and CPU operations per Pixel calculated for different degrees of FF-Model
The memory calculation is based on the image resolution 1280x960 and FF-Model resolution of 4.

The accuracy parameters for a perspective/central-projection camera are available in Table 7.1. The “RMSE” and “Max Error” values are calculated over the entire image using all of the observed control points. As we can see the free-function model delivers a good calibration accuracy with smaller RMSE and maximum error. By considering the amount of required memory, CPU operation per pixel and the accuracy values, we can select the optimal degree for the free-function modelling for a considered use-case. We should also mention that as previously discussed in [4.4.1], there is no risk of correlation between intrinsic and extrinsic parameters as an initial calibration model would be first estimated to avoid overestimation.

In next chapter, similar to the process performed here, the free-function model would be used for an omnidirectional camera and the results will be analysed.

7.2. Performing Free-Function Calibration on Omnidirectional Camera

This process is very similar to the one discussed in [7]. In this chapter we use the free-function calibration model introduced in [4] to model the lens distortion of an omnidirectional camera with the specifications in [6.1] using the test stand calibration method from [6.6]. We calibrate the camera initially using the extended omnidirectional model [3.7.1] [3.7.2] [3.7.3].based C. Mei model [2.4.2.5]. Then the primary and local lens distortion-field using Gaussian interpolation [4.3] are generated.

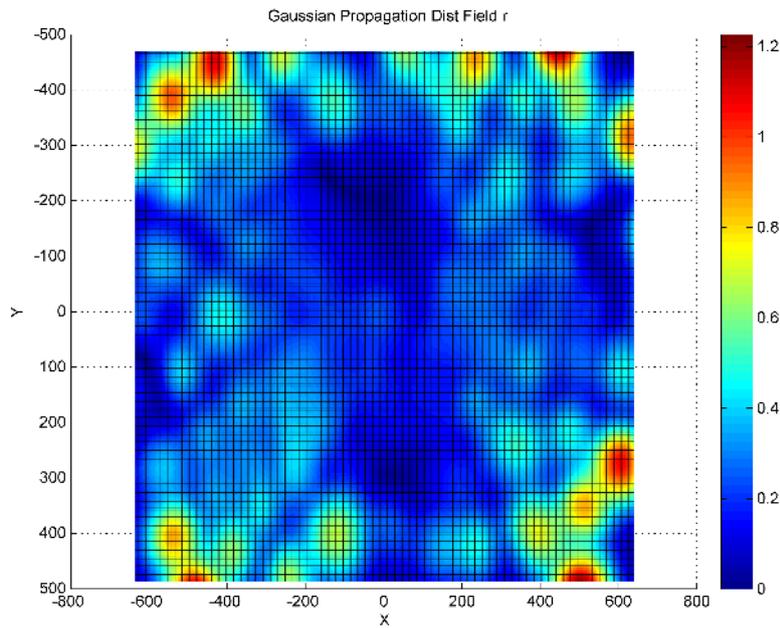
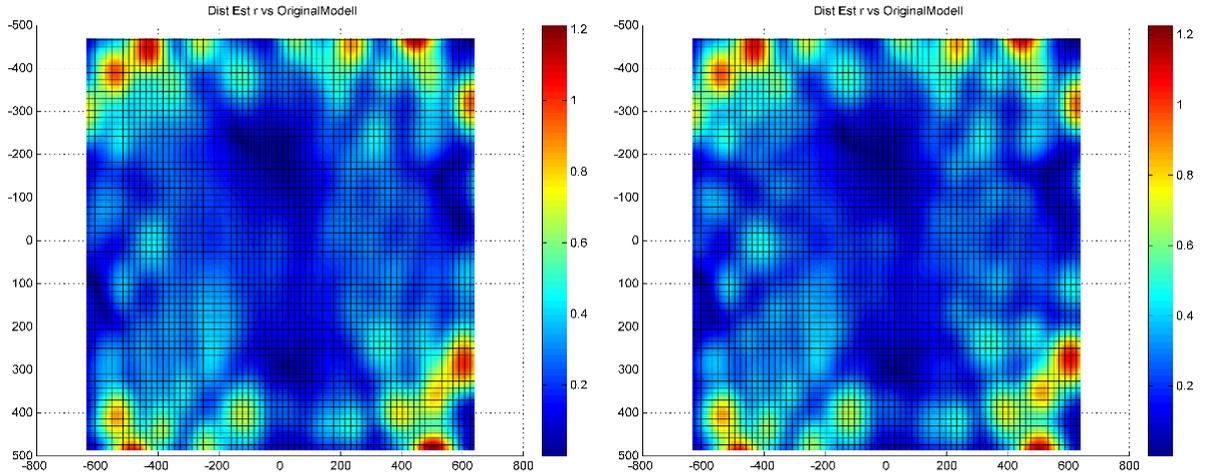


Fig. 7.5. Example of measured local distortion field (input data) generated using Gaussian-interpolation for a fisheye image
Distribution domain size= 300, StdDev factor= $\frac{1}{5}$, The colour-bar and axes coordinates are in pixel units

(Courtesy of Robert Bosch GmbH)

Fig. 7.5 shows the amount of the local lens distortion in both x and y directions for a fisheye camera. Having these data as input and adding them to the distortion-field generated by the original calibration model, we can perform the free-function model estimation. After that, by subtracting the distortion-field generated by the free-function model from the one generated by original model, we can see how good the local distortions are modelled.

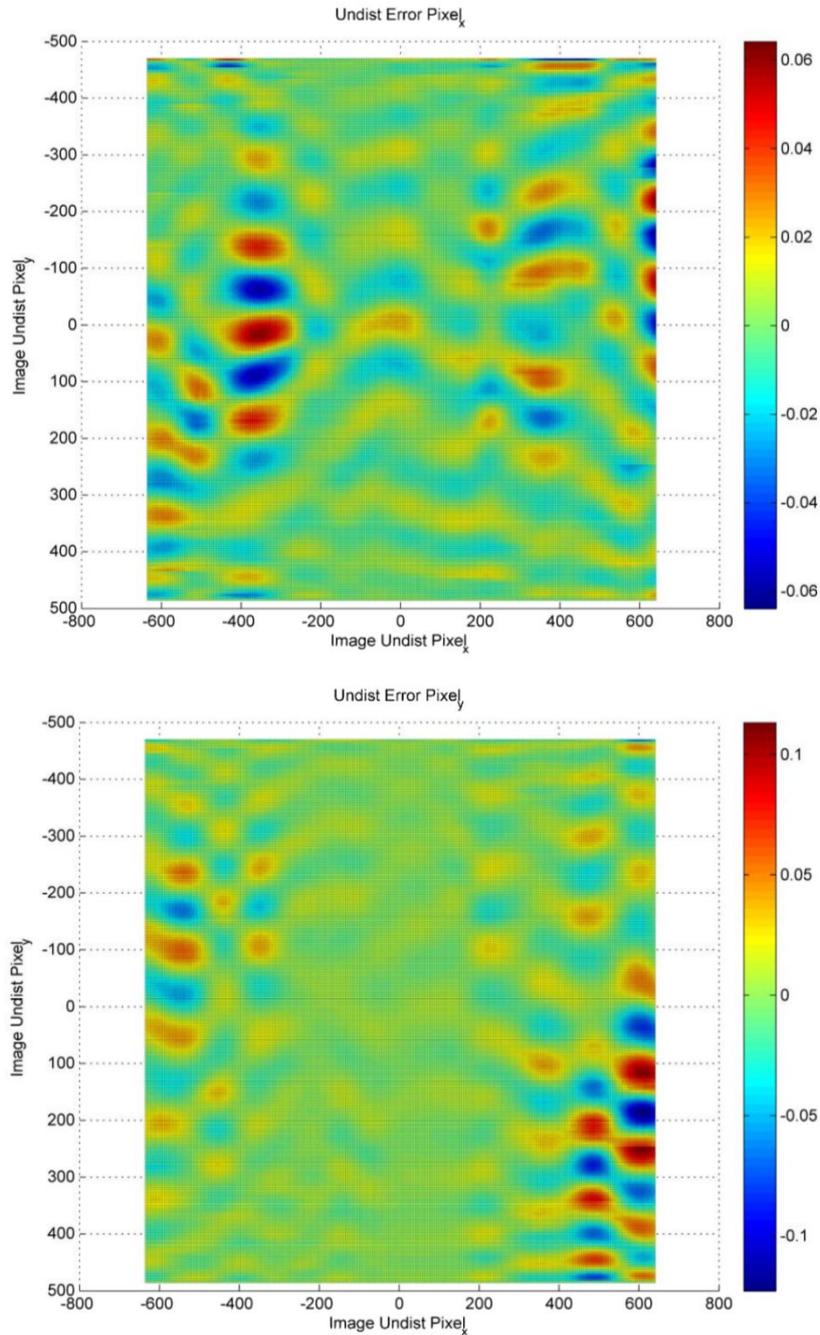


a b Fig. 7.6. Result of free-function distortion modelling (of local lens distortion) for a fisheye image (in radial direction), Distribution domain size= 300, StdDev factor= $\frac{1}{5}$
 a: free-function degree= 16 b: free-function degree= 32
 The colour-bar and axes coordinates are in pixel units
 (Courtesy of Robert Bosch GmbH)

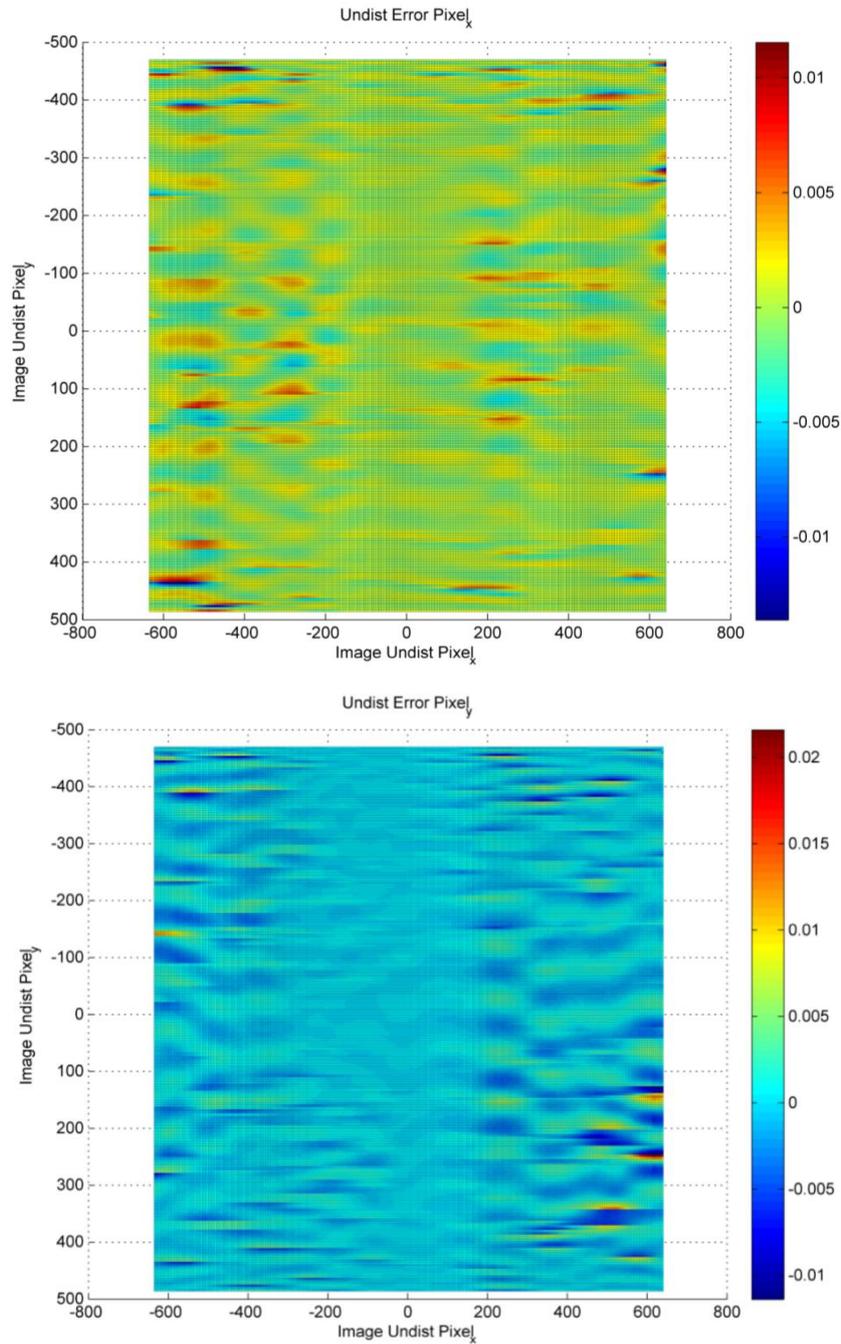
We performed distortion modelling using the free-function model of degree 16 and 32 ($n=16$ and $n=32$) using the distortion-field in Fig. 7.5 and the original model as input data. The results of local lens distortion modelling is illustrated in Fig. 7.6. The calibration is performed following the recommended process of autonomous calibration by S. Abraham (Abraham & Hau, 1997) (Abraham, 2004). As we have also seen in [7] the quality of local lens modelling increases by using higher degrees for the free-function model. The accuracy measures are available in the next chapter.

7.2.1. Accuracy and Error Analysis

In this chapter we study the accuracy parameters of the free-function model with the degree of 16 and 32 ($n=16$ and $n=32$) and compare them to the accuracy parameters of a standard calibration model. Here we analyse the accuracy of free-function using the input data from [7.2] and Fig. 7.5 for a fisheye camera. After modelling the distortion, the residuals or the difference between the input distortion field and the modelled one are calculated. Using them, we estimated how good the free-function model could practically model the distortion field.



a Fig. 7.7. The remaining errors from free-function calibration for an omnidirectional image
 b Using free-function of degree 16, a: errors in image x direction, b: errors in image y direction



a Fig. 7.8. The remaining errors from free-function calibration for an omnidirectional image
 b Using free-function of degree 32, a: errors in image x direction, b: errors in image y direction

Fig. 7.8 shows the remaining errors of local lens distortion modelling using the free-function model of degree 16 and 32 on the input data from Fig. 7.5. We can see that the errors are very small which means the local distortion is very well modelled.

Parameter Name	Extended Omnidirectional (9 Parameter)	free-function Model (degree 16)	free-function Model (degree 32)
RMSE _x	0.199	0.012	0.001
RMSE _y	0.188	0.016	0.001
Max Error x	0.531	0.061	0.014
Max Error y	0.581	0.123	0.022

Table 7.3. Accuracy analysis of the free-function model in comparison with a standard distortion model
(All the parameters are in pixel units)

In Table 7.3, the calculated accuracy parameters after performing camera calibration are available. The “RMSE” values are calculated using all of the observed control points in image, the “Max Error” values are calculated over the entire image and we can see the accuracy of free-function model compared to other distortion models.

7.3. Further Analysis of Lens Distortion Modelling using FF-Model on Real-World Data

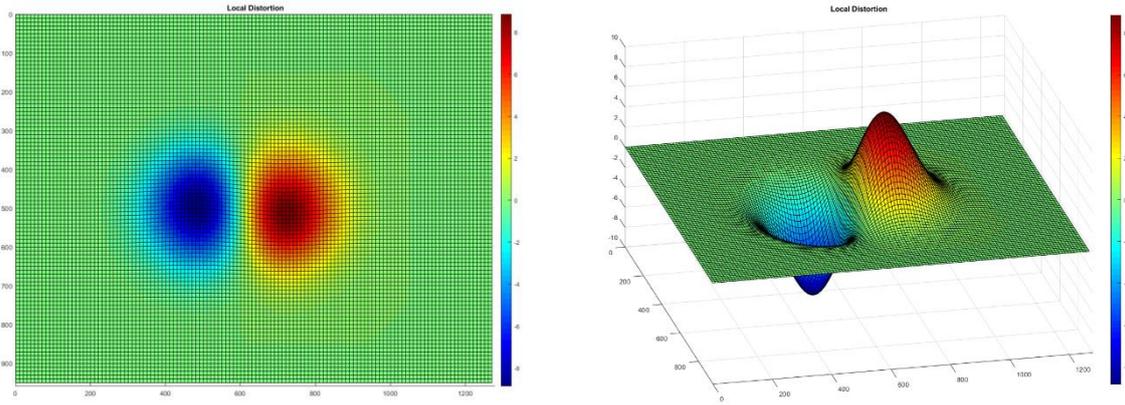
Having studied the concept of local lens distortion modelling in [4.3] and the calibration process-flow in [4.4], and having performed the distortion modelling in 7.1.1[7.1] and [7.2], we can now continue with a further in-depth analysis of the free-function model using real-world images to study the effectiveness of the local lens distortion modelling on real use-cases. For this purpose, we have used a stereo front-view camera illustrated in Fig. 2.9b and Fig. 6.1b, taking stereo pair images from a vehicle placed in front of the camera (Fig. 7.9) to generate disparity images or disparity maps. As illustrated in Fig. 2.5b and Fig. 2.12b, the disparity maps could be used to generate 3D or depth information from stereo images. In order to create a visually measurable scenario, a large local distortion was added to the left images of the stereo camera. This local distortion was added to the images during the camera calibration phase and also during image recording of the front vehicle. In this test, we review the distortion modelling and the quality of the generated disparity map to evaluate the effectiveness using free-function model, and compare the results with another conventional model [2.4.3.6].



a	b
---	---

 Fig. 7.9. Stereo image pairs taken from a front vehicle using a stereo front-view camera
The distance of the vehicle to the camera is around 10 [m]
a: left image, b: right image (Courtesy of Robert Bosch GmbH)

The added local distortion is illustrated in Fig. 7.10. This distortion is in the form of a pair of smooth cone-like functions also known as blobs. These distortion irregularities are created based on a function of the third power of the radial distance from its centre, with an effective size of around 350 [Pixel] and amplitude of around 8 [Pixel]. This amplitude and size are selected purely for easier visual differentiation in our test and are arbitrary values if they represent possible and measurable optical distortion. The density of a disparity map is usually more sensitive to errors in y direction, especially if one-dimensional image matching solutions are used, which have been developed excessively in the past decade (Szeliski, 2010). Therefore, a substantial error in y image coordinates could lead to a decrease in the density of stereo matching in corresponding rows of the rectified images. As this reduction in density is also visually identifiable we added some local distortion to the left image in the vertical (y) direction. We know that the free-function modelling is performed in both directions for both x and y coordinates. Thus the quality of distortion modelling in x and y direction would be very comparable to each other if the local distortion is present in both directions. Nevertheless, adding the distortion in the y direction creates a comparable scenario when we are comparing calibration results and the quality of their generated disparity maps.



a b Fig. 7.10. Local distortion in the form of a pair of blobs, a: 2D view, b: 3D view



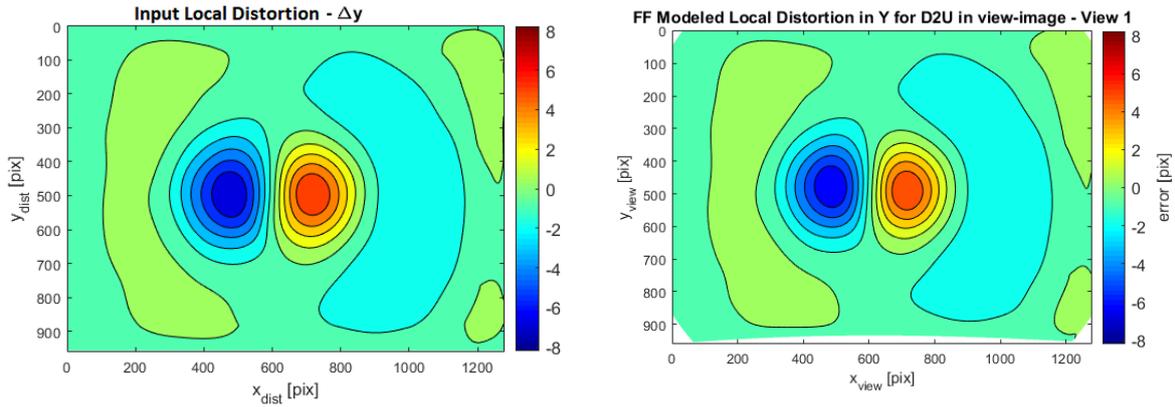
Fig. 7.11. Left image of the stereo pair after adding the local distortion.

This local distortion was added to the left image of the stereo pair as illustrated in Fig. 7.11. As mentioned previously, this local distortion was also available during the camera calibration phase at the test-stand. First, we calibrated the stereo camera using Tschebycheff calibration model of third order [2.4.3.6] as a conventional model. After that we used the free-function model of degree 16 as the calibration model using the results of the calibration with Tschebycheff model as initial data as described in [4.4.4].

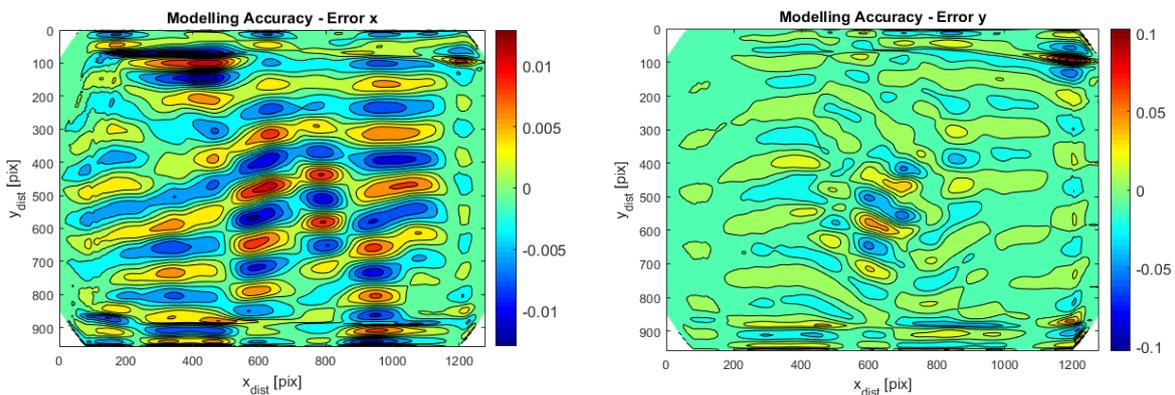
Calibration Model	RMSE x	RMSE y	Max Error x	Max Error y
Tschebycheff 3 rd order (18 parameter)	0.151	2.272	0.416	8.202
FF-Model (Degree 16)	0.002	0.016	0.013	0.103

Table 7.4. Accuracy analysis of the Tschebycheff model and the FF-Model on an image with large local distortion (All the parameters are in pixel units)

In Table 7.4, the result of the calibration using both models are available. Compared to a camera without this local distortion (Table 7.1), the Tschebycheff model was able to model a small part of the local distortion by adapting the equivalent focal length parameter and compensating the effects with its distortion parameters (Fig. 7.12a, Input of the FF-Model). However, the remaining errors are still very large. Nevertheless, we can see that the remaining local distortion in y direction could be effectively modelled using the free-function model. It is even possible to predict the difference in the density of the resulting disparity maps of the two models by having the above calibration residuals. Now we continue by performing further accuracy analysis of the free-function model on the left image of the stereo pair.

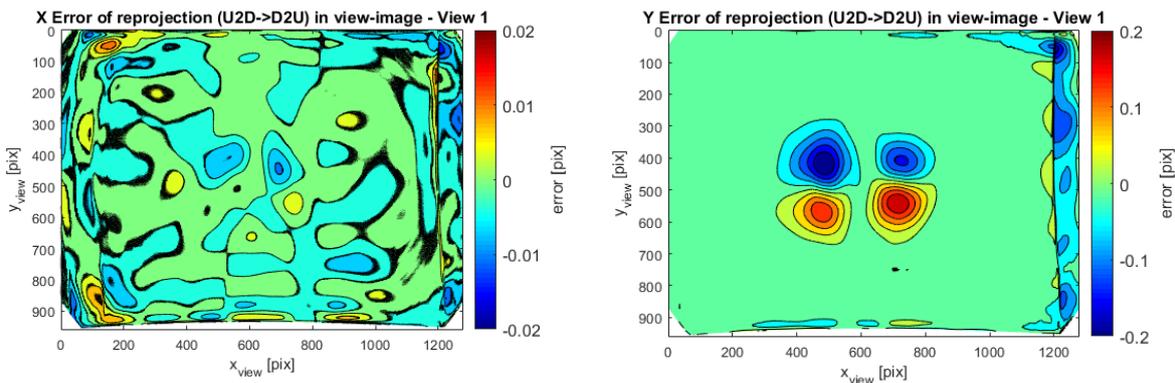


a **b** Fig. 7.12. Modelling accuracy of the FF-model (the initial measured local distortion is prepared in the form of a residual map as input data for the FF-modelling) calculated for the left image of the stereo pair, a: input local distortion in y direction, b: modelled local distortion in y direction



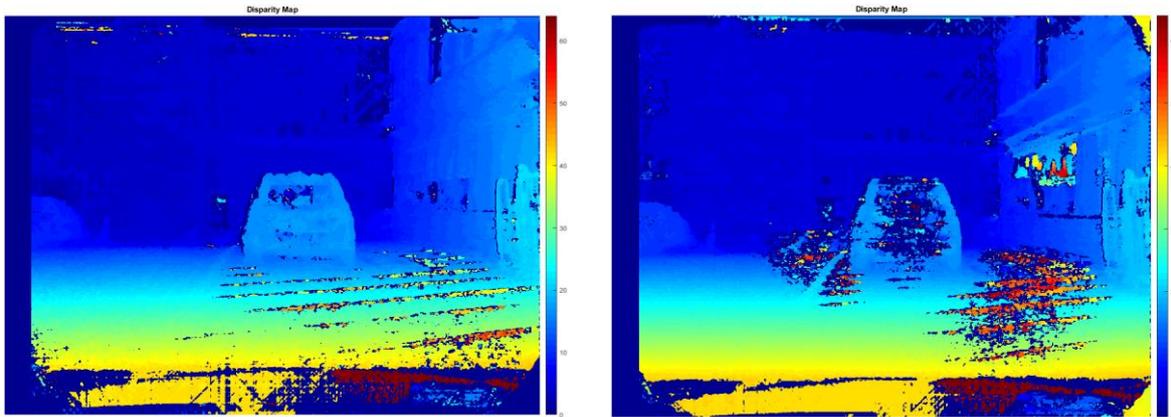
a **b** Fig. 7.13. Modelling accuracy of the FF-model (residuals) in distorted image which is the undistortion function or D2U and transforms the distorted to undistorted coordinates, a: x coordinates, b: y coordinates

We can see in Fig. 7.12 and Fig. 7.13 that the local distortion has been modelled very well. The modelling accuracy in x direction is higher because the local distortion is only available in y direction. Calculating the inverse of the undistortion function, we will get the distortion function (U2D) leading to similar accuracies with maximum error in x and y equal to 0.014 and 0.102 [Pixel] respectively. To get the final accuracy results, a closed loop coordinate calculation is done, where undistorted coordinates are transformed via the distortion function (U2D) into distorted coordinates and then transformed back via the undistortion function (D2U) to undistorted image coordinates. The error of this closed-loop calculation could be considered as the total possible modelling error, which includes distortion inversion.



a **b** Fig. 7.14. Total modelling accuracy of the FF-model (residuals) for left image of the stereo pair in a closed-loop calculation, a: x coordinates, b: y coordinates

As we can see in Fig. 7.14, the total modelling error of the free-function model measured in a closed-loop calculation is around 2 times the modelling error in each distortion or undistortion transformation and is still very acceptable even if we have very large local distortions in image, which is the case here. Now we proceed forward with generating the disparity maps of the stereo image pair using the calibration results of both models for the stereo camera, and compare the quality and the density of the disparity map.



a	b
---	---

 Fig. 7.15. Disparity maps generated using calibration results for the stereo camera.
a: using the FF-model (degree 16), b: using the Tschebycheff Model (3rd order, 18 parameter)
The disparity map is generated using the disparity function in MATLAB 2018b (9.5.0.944444) 64-bit

Fig. 7.15 shows the difference between the densities of the disparity maps, generated, using the results of the stereo camera calibration utilizing both calibration models. As it is visible, using the free-function model, the disparity map is available in much higher densities for the front vehicle and its surroundings especially at image areas where the local distortion was present. This implies that the distortion modelling has been much more effective using the free-function model and this model was able to deliver high accuracies even with the large added local distortion in Fig. 7.10, as this was verified in details in this chapter. In the next chapter, we further discuss the accuracy parameters and provide some interpretations and thoughts on the accuracy of the free-function model.

7.4. Discussion on Accuracy of Free-Function Model

We tested the camera calibration using the free-function model in [7.1] and [7.2], and performed error analysis in [7.1.1] and [7.2.1] and have seen respectable accuracy results. The calibration processes is based on the process-flow described in [4.4] and illustrated in Fig. 4.5. This process employs the free-function model as a posteriori calibration step, getting the input from the preliminary modelled distortion field and the local distortion field derived from the calibration residuals in image. The relatively small residuals remaining after the free-function calibration represent the high flexibility of this model to fit itself to different distortion fields with localized non-symmetrical optical behaviour. It means that, if applied with enough degrees, which is the degree used for the expanded multivariate polynomials utilized in this model, it will fit to almost any conventional distortion model with local anomalies with good quality.

The above discussion implies that the input information has to be dealt with extra caution to avoid any big or systematic errors in them, as they will be modelled alongside the actual camera distortion. For example, the test stand is measured with micrometres accuracy threshold discussed in [6.2.4]. Generally, any error in the test-stand 3D coordinates can potentially be transferred to the residuals after the preliminary calibration step. The local lens distortion field [4.3] is the interpolation of the residuals measured on dense and homogeneously distributed points. Thus, in this case, random errors in the test stand coordinates will not be modelled by the free-function calibration and are eliminated in the interpolation phase, as it is apparent that the interpolation or average of a random field with small numbers and high parameter count be statistically considered equivalent to zero. However, if there are any existing local systematic errors in the three-dimensional coordinates of the test stand, although they are below the accuracy threshold and will not harm the calibration's targeted accuracy, they will be present – to some extent - in the interpolated residuals and could be modelled in the a posteriori calibration.

In order to avoid the problem of systematic errors in the 3D coordinates, we recommend calibrating the test stand and measuring the three dimensional point coordinates at least three times separately and calculate the mean values of the coordinates to avoid having systematic errors. This will ensure that the coordinates of the test stand are free of systematic errors and enables us to correctly consider the interpolated residuals to be local lens distortion, which has been discussed in [4.3]. By considering this recommendation and taking this measure for any single-shot test stand calibration, we can confidently consider the results of the free-function calibration as the lens distortion with modelled local lens behaviours with respectable accuracy.

8. Summary and Conclusion

In this study, we reviewed the current techniques and methods in photogrammetry - especially close-range photogrammetry – and focused on camera calibration. We reviewed the new evolving field of video-based driver-assistant systems, their requirements and their applications. Then in order to accurately calibrate variety of perspective and fisheye cameras that are potentially used in driver-assistant systems (and also many other fields of photogrammetry), and later use the calibration data effectively in our applications, we reviewed many calibration models and also the existing suitable calibration hardware for this purpose.

Exclusively of fisheye cameras and a general omnidirectional projection, we used a decent camera calibration model from Christopher Mei and extended it further to address our needs and functionality requirements. These extensions enable us to use the camera calibration model in special mobile hardware and embedded systems with low processing power and near-run-time characteristics. The extended model has also many advantage over other calibration models such as flexibility, and physical interpretation capability.

We also introduced the free-function model as a flexible and advantageous model for camera distortion modelling. Free-function calibration is a new approach for modelling the overall image distortion together with the local lens distortions that are estimated using a standard model during the calibration process. As we have shown, using free-function model on different lens designs, one can achieve good calibration accuracies by modelling the very local lens distortion taking benefit from the flexibility of this model. This feature is achieved by considering unknown functions in the distortion model and estimating them numerically. Furthermore, we introduced optimization strategies for recalculation and image rectification. These optimizations are also used to minimize the amount of required device memory. We have also introduced further possibilities of enhancing the speed and reducing the required memory, which could be advantageous considering the existing hardware specifications and accuracy requirements. (Nekouei Shahraki & Haala, 2015 A)

The free-function model has very good flexibility, which means by changing the degree of model, we can achieve the necessary calibration accuracy. The simplicity of implementation of this model for recalculation and image rectification (i.e. linear vector-based multiplication) brings many advantages to variety of computational platforms such as FPGAs, x86 and ARM processors, and makes it possible to benefit from variety of parallel-processing techniques. This model is capable of being used in runtime even on processing platforms and embedded systems with limited processing power or memory resources, and is an ideal calibration model for using in variety of machine vision solutions.

We also discussed several important requirements for accurate camera calibration that we later used in hardware test stand design phase. We designed and developed two different test stands in order to realize the specifications and geometrical features of multiple-view test-field-based camera calibration referred to as bundle-block calibration. One of their special geometrical characteristics is the uniform point distribution, which corresponds to the uniform motion. This uniform and dense point distribution also helps us to have extrapolation capabilities in the calibration model further beyond the last detected image point at the image edge. Furthermore, such a point distribution is beneficial when using calibration models such as free-function model that enable us to model of local lens distortion with good accuracy

and quality all over the image. A very important feature of this test stand is having the capability of performing camera/sensor alignment testing, a feature which is very important for testing the geometrical alignment of the internal mechanical elements of each camera. Using automated machines and algorithms in test stand calibration increased the stability and accuracy of the calibration and thus ensured the quality and speed of the calibration for cameras used in video-based driver assistant systems. (Nekouei Shahraki & Haala, 2015 B). These test stands are capable of performing automatic camera calibration, suitable for applications such as series-production of cameras.

As an accuracy –and flexibility evaluation step for the free-function model, we tested the free-function calibration model on real-world data using a stereo camera with added large local distortions taking images from a front vehicle similar to the conditions where real-world use-cases are defined. By performing the camera calibration, we compared the calibration results and accuracy parameters of the free-function model to a conventional calibration model. Using these calibration results, we generated a set of disparity maps and compared their density and availability, especially on the areas where the local distortion was present. We used this test to compare the capabilities of the proposed model to conventional ones in real-world situations where large optical distortions could be present that cannot be easily modelled with conventional calibration models.

The higher modelling capability and accuracy of the free-function model will generally influence those functions that are using the information of the disparity map or the derived 3D information as part of their input data and potentially leads to the better functionality or even their availability if local distortions are present in the image. We can name few functions used in driver-assistant systems that could potentially use and benefit from stereo 3D information such as stereo-based emergency braking, collision avoidance, manoeuvre assistant, pedestrian protection, adaptive cruise control, platooning and many more that have been discussed previously in [2.3]. The benefits of this model are not limited to driver-assistant systems. There are many more use-cases in photogrammetry and computer-vision where a higher calibration accuracy is beneficial on hardware such as low-cost optics where sometimes optical distortion are available that cannot easily be modelled with classical models. These use-cases could all benefit from the flexibility and modelling accuracy of the free-function model.

9. Outlook

In this work, we have studied the repeatability accuracy of the camera calibration using the newly designed test stands. The design of the test stand will enable us to expose the camera to different environments such as different temperature ranges, humidity levels, etc. and potentially enables us to model the camera's optical behaviour under different conditions. This is especially helpful in applications in which the camera experiences different environmental conditions and in which an in-situ re-calibration is not accurately possible or practical. Thus, it could be expected to see use-cases in which these test stands, or similar ones based on the design-criterion are used for environmental modelling of special optical systems.

We have also studied the calibration accuracy of the introduced free-function for different types of cameras using the newly designed test stands which shows the flexibility of this calibration model that help us model complex optical behaviours such as local distortions. Therefore, it could be confidently said that there are yet some use-case potentials waiting to be discovered. There has recently been an increasing interest in different fields to benefit from free-form optical systems that deviate from classical symmetric optical designs, and there would very likely be a need for an accurate optical modelling. This can doubly encourage us in the future to test this calibration model on different optical systems which are based on free-form optical designs.

Furthermore, the free-function calibration model has the advantage of modelling optical distortions that comprise of overlaying distortion fields. In case those distortion fields are all modelled with the free-function model, then those models and their estimated parameters could be added to each other to form the final distortion model. This could potentially enable us to measure a complex optical system by modelling its single optical elements and finally merging the calibration models together. This could potentially turn out to be advantageous in case a complex optical system is difficult to be measured as a whole, but possible to be measured as single elements or modules.

10. Appendix: Statistical Model and Error Analysis

In this chapter, we have a short overview of the statistical model and the estimation techniques used to calculate the parameters and accuracy measures.

10.1. Parameter Estimation System and Algorithms

The statistical model that we use is the famous “Least Squares Adjustment” (Gauss-Markov Linear model). We use this model when the number of observations is bigger than the number of unknowns and therefore there is no unique solution. Here we discuss a brief summary of the least squares method. For more details, please refer to reference books (Amini, 2006) (Luhmann, et al., 2011). If we want to solve an equation system, first the observations and unknowns have to be parameterized (Eq. 10.1):

$$\mathbf{L} = (L_1, L_2, \dots, L_n)^T = \text{observations vector}$$

$$\mathbf{X} = (X_1, X_2, \dots, X_u)^T = \text{vector of unknowns}$$

$$\tilde{\mathbf{L}} = \varphi(\tilde{\mathbf{X}}) = \begin{bmatrix} \varphi_1(\tilde{\mathbf{X}}) \\ \varphi_2(\tilde{\mathbf{X}}) \\ \dots \\ \varphi_3(\tilde{\mathbf{X}}) \end{bmatrix} = \text{functional model} \quad \text{Eq. 10.1}$$

The observation vector is replaced by vector $\tilde{\mathbf{L}}$ since the real values are unknown (\mathbf{X}^0 is the approximate value for the unknown $\tilde{\mathbf{X}}$). Next, the variables are initialized:

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v} = \varphi(\hat{\mathbf{X}})$$

$$\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}} \quad \text{Eq. 10.2}$$

$$\mathbf{L}^0 = \varphi(\mathbf{X}^0)$$

Using a Taylor series, we can linearize the functional model:

$$\mathbf{L} + \mathbf{v} = \varphi(\mathbf{X}^0) + \left(\frac{\partial \varphi_1(\mathbf{X})}{\partial X_1} \right) (\hat{\mathbf{X}} - \mathbf{X}^0) = \mathbf{L}^0 + \left(\frac{\partial \varphi_1(\mathbf{X})}{\partial X_1} \right) (\hat{\mathbf{x}}) \quad \text{Eq. 10.3}$$

With the linearized model, the Jacobian matrix is introduced:

$$\mathbf{A}_{n,u} = \left(\frac{\partial \varphi}{\partial \mathbf{X}} \right)_0 = \begin{bmatrix} \left(\frac{\partial \varphi_1(\mathbf{X})}{\partial X_1} \right)_0 & \left(\frac{\partial \varphi_1(\mathbf{X})}{\partial X_2} \right)_0 & \dots & \left(\frac{\partial \varphi_1(\mathbf{X})}{\partial X_n} \right)_0 \\ \left(\frac{\partial \varphi_2(\mathbf{X})}{\partial X_1} \right)_0 & \left(\frac{\partial \varphi_2(\mathbf{X})}{\partial X_2} \right)_0 & \dots & \left(\frac{\partial \varphi_2(\mathbf{X})}{\partial X_n} \right)_0 \\ \dots & \dots & \dots & \dots \\ \left(\frac{\partial \varphi_n(\mathbf{X})}{\partial X_1} \right)_0 & \left(\frac{\partial \varphi_n(\mathbf{X})}{\partial X_2} \right)_0 & \dots & \left(\frac{\partial \varphi_n(\mathbf{X})}{\partial X_n} \right)_0 \end{bmatrix} \quad \text{Eq. 10.4}$$

Any observation has a standard deviation which corresponds to its weight in the adjustment and the weight matrix K is form:

$$K_{ll} = \begin{bmatrix} \sigma_1^2 & p_{12}\sigma_1\sigma_2 & \dots & p_{1n}\sigma_1\sigma_n \\ p_{12}\sigma_2\sigma_1 & \sigma_2^2 & \dots & p_{2n}\sigma_2\sigma_n \\ \dots & \dots & \dots & \dots \\ p_{n1}\sigma_n\sigma_1 & p_{n2}\sigma_n\sigma_2 & \dots & \sigma_n^2 \end{bmatrix} \quad \text{Eq. 10.5}$$

After evaluating the Jacobian matrix, we can perform the least squares adjustment by continuing to write the functional model, the stochastic model, and the matrix of normal equations:

$$\hat{l}_{n,l} = l_{n,l} + \hat{v}_{n,l} = A_{n,u}\hat{x}_{u,l} \quad \text{Eq. 10.6}$$

$$Q_{ll} = \frac{1}{s_0^2} K_{ll} = P^{-1} \quad \text{Eq. 10.7}$$

$$N_{u,u} = A_{u,n}^T P_{n,n} A_{n,u} \quad \text{Eq. 10.8}$$

Now the vector of unknowns is estimated and the observations vector is updated:

$$\begin{aligned} n_{u,l} &= A_{u,n}^T P_{n,n} l_{n,l} \\ \hat{x}_{u,l} &= Q_{u,u} n_{u,l} \\ \hat{l}_{n,l} &= A_{n,u} \hat{x}_{u,l} \end{aligned} \quad \text{Eq. 10.9}$$

Next, the standard deviation a posteriori and the variance-covariance matrix could be estimated:

$$\begin{aligned} \hat{s}_0 &= \sqrt{\frac{v^t p v}{n - u}} \\ K_{u,u} &= \hat{s}_0^2 Q_{u,u} \end{aligned} \quad \text{Eq. 10.10}$$

In an iterative adjustment, the updated/estimated variables are used again as initial values to get better estimations and the iteration stops when the corrections are relatively small.

10.2. Confidence Interval Test

The confidence interval test (Luhmann, et al., 2011) is about how good our observations are distributed corresponding to an ideal Gaussian probability distribution function for infinite independent observations of a same variable where, σ is the standard deviation and μ is the expectation.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right) \quad \text{Eq. 10.11}$$

On the other hand, limited numbers of observations have the t-student distribution other than a Gaussian distribution. The t-student test is described in Eq. 10.12:

$$P\{C_u \leq \hat{x} \leq C_o\} = 1 - \alpha$$

$$P\{\hat{x} < C_u\} = \frac{\alpha}{2}$$

$$C_u = \hat{x} - t_{f,1-\frac{\alpha}{2}} S_{\hat{x}} \quad \text{Eq. 10.12}$$

$$C_o = \hat{x} + t_{f,1-\frac{\alpha}{2}} S_{\hat{x}}$$

$$C_o - C_u = 2 t_{f,1-\frac{\alpha}{2}} S_{\hat{x}} = \text{Observations Interval}$$

When we perform the stochastic t-student test having a null hypothesis H_0 , the degree of freedom (f), the standard deviation, and the observations interval, we get the confidence interval of the observations in percentage. The null hypothesis can be set to if our observations are having a certain confidence interval. The outcome of the test is a factor that represents how well the observations were performed, including no outlier gross errors, and it shows in percentage that how confident we are about the observations.

10.3. Correlation, Independency and Significance

In any adjustment operation, after the estimated covariance matrix is obtained, we can calculate a measure for correlation between the estimated unknown parameters. This is described in Eq. 10.13.

$$K_{\hat{x}\hat{x}} = \hat{S}_0^2 Q_{\hat{x}\hat{x}} \begin{bmatrix} \hat{S}_1^2 & p_{12}\hat{S}_1\hat{S}_2 & \dots & p_{12}\hat{S}_1\hat{S}_2 \\ p_{12}\hat{S}_2\hat{S}_1 & \hat{S}_2^2 & \dots & p_{2u}\hat{S}_2\hat{S}_u \\ \dots & \dots & \dots & \dots \\ p_{u1}\hat{S}_1\hat{S}_1 & p_{u2}\hat{S}_1\hat{S}_2 & \dots & \hat{S}_u^2 \end{bmatrix} \quad \text{Eq. 10.13}$$

$$p_{ij} = \frac{\hat{S}_{ij}}{\hat{S}_i\hat{S}_j} \quad -1 \leq p_{ij} \leq +1$$

This test searches for those parameters, which were not estimated independently and shows how good our adjustment model is fitting to the reality (observed data).

Parameter Significance is defined based on its value (amplitude) and its precision or the estimated standard deviation. An estimated parameter l_i is significant, when its estimation precision C_i is much smaller than its absolute value $|l_i|$, which implies $S_i \gg 1$.

$$S_i = \frac{|l_i|}{C_i} \quad , \quad S_i \gg 1 \quad \text{Eq. 10.14}$$

In other words, if the accuracy is smaller than the parameter itself, it is conclude that we have control over that parameter and if the estimation precision is close to the absolute value $S_i \approx 1$ then there is small or no control over that parameter and it is insignificant.

11. Appendix: Pre-Analysis and Network Design

Pre-analysis is referred to the operation which is performed while designing an (adjustment) network and helps us to decide how the network should geometrically be designed and how accurate the observations and measurements have to be to ensure the required accuracy and significance for the unknown parameters (Grafarend & Sanso, 1985) (Kuang, 1996) (Amiri-Simkooei, et al., 2012).

Typically, pre-analysis (as it is clear from its name) is performed during the design of the network and is classically divided in different steps, which are called “zero, first, second, and third order design”. (Grafarend, 1974) (Amiri-Seemkooei, 2001). Here we are mostly interested in the second order design. As we know, the accuracy of the observations is affecting the accuracy of estimated parameters and the cost of the project and these are the two most important factors in any adjustment operation.

As illustrated in Fig. 11.1, when we have an approximate design of the network (design matrix A), and some expectations for the standard deviation of unknowns ($K_{u,u}$), using those expectations we can calculate the required accuracy of the observations (P_{nn}) to determine accuracy thresholds (with a certain factor of safety). This means that (later in the measurement phase) the accuracy of the observation should be at least within those limits (or higher than the determined thresholds) so that the accuracy of estimated parameters meets our expectations.



Fig. 11.1. Pre-Analysis process flow

12. Appendix: Post-Analysis and Network Design Evaluation

If the initial design is available, we can perform the network pre-analysis in different steps or iteratively, and improve different aspects of a network at each iteration. Here we define post-analysis as a later iteration of pre-analysis, which is performed in a posteriori step at a final stage of network design improvement. For example, after performing pre-analysis of zero, first and second order, and evaluating the necessary accuracy of the observations and measurement and after performing network adjustment, we can perform the network pre-analysis third-order design (Grafarend, 1974) and consecutively revise the network geometry in a first-order design optimization. This optimization could be performed manually or analytically (Amiri-Seemkoei, 2001) (Cross, 1985). Therefore, this helps us optimize the network configuration and gives us useful information about the accuracies. This is also useful for making decisions over the estimated parameters. The results of post-analysis are later used for updating the network geometrical design or improving the observation accuracy if necessary.

As discussed above, the implementation and strategy of post processing, especially for camera calibration purposes, could be to perform the third order design after performing the zero, first and second order design. Typically, after marker-detection is done and the system of equations was set up, we have the initial design of the network and also the expectations (expected limits) of the standard deviation of unknowns. Therefore, we can perform post-analysis to re-evaluate the geometrical design of the test stand, and verify if the network geometrical configuration and accuracy of the test stand calibration and image measurement are really sufficient for satisfying our required accuracy of camera intrinsic parameters.

Similar to the first pre-analysis optimization steps, after performing post-analysis, we get thresholds for the accuracy of each target point (e.g. assuming they are measured with the same accuracy and also having a factor of safety equals to 3 i.e. 99.7% reliability). If the accuracy of test stand point measurement is higher than this threshold, this would ensure achieving those expected standard deviation of the estimated (intrinsic) parameters. If after post analysis we witness any geometrical deficiency in the test stand such as parameter insignificance (in test stand target points) or high correlation between independent parameter, the geometrical design of the test stand should be improved. As discussed in this chapter, this is achieved in a posteriori third order, and an iterative first-order design optimization.

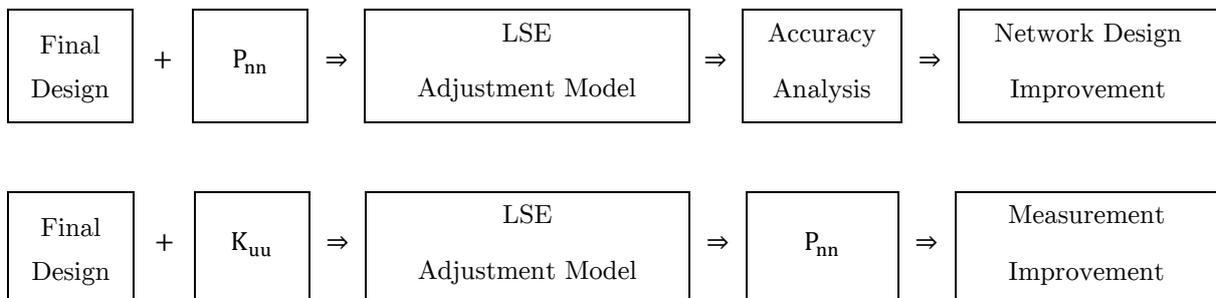


Fig. 12.1. Post-Analysis process flow

13. Appendix: Evaluation of Mathematical Models for Analysing Fisheye Distortion

In this appendix, we provide the information regarding the results of different models for calibrating several fisheye cameras. Therefore, we can evaluate their accuracy, and compare the results. This evaluation has been done by (Abraham, 2012) similar to the validation of models by (Schneider, et al., 2009) and gives us a good overview on how suitable different calibration models are for calibrating cameras with different optical designs.

We have a set of four different fisheye cameras. As we can see in Table 13.1, varieties of calibration models have been selected.

Index	Model Name	Parameter Count	Model Description
1	Equidistance	1	$r' = c \cdot \Phi$
2	Stereographic	1	$r' = c \cdot \tan\left(\frac{\Phi}{2}\right)$
3	Modified Orthogonal	2	$r' = c \cdot \sin(a \cdot \Phi)$
4	Extended Modified Orthogonal	3	$r' = c \cdot \sin(a \cdot \Phi) + P2(\Phi)$
5	Extended Modified Orthogonal	4	$r' = c \cdot \sin(a \cdot \Phi) + P2(\Phi)$
6	Scaramuzza	5	
7	Radial Polynomial	5	$r' = P5(\Phi)$
8	Tschebycheff Polynomial	5	$r' = T5(\Phi)$
9	Christopher Mei	5	

Table 13.1. Different models used for calibrating a set of fisheye cameras

The fisheye cameras used in this test all have a HFOV bigger than 180 degree as visible in Fig. 13.1:

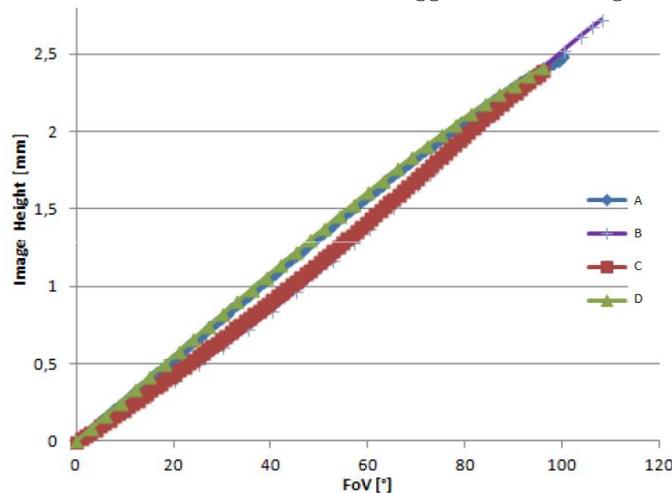


Fig. 13.1. Angular projection figures of the four fisheye test-cameras (A..D)

The process of calibration and accuracy analysis is performed in three steps:

1. Performing Least-Squares Calibration using the available design data
2. Illustration of the deviation between the modelled projection and design data
3. Calculation of maximum error between the mathematical model and design curve

After performing the camera calibration using the design data, the results of calibration from all cameras are brought together. The results are given separately for each model in the following figures. The curves corresponding to the each fisheye lens are illustrated with A, B, C and D:

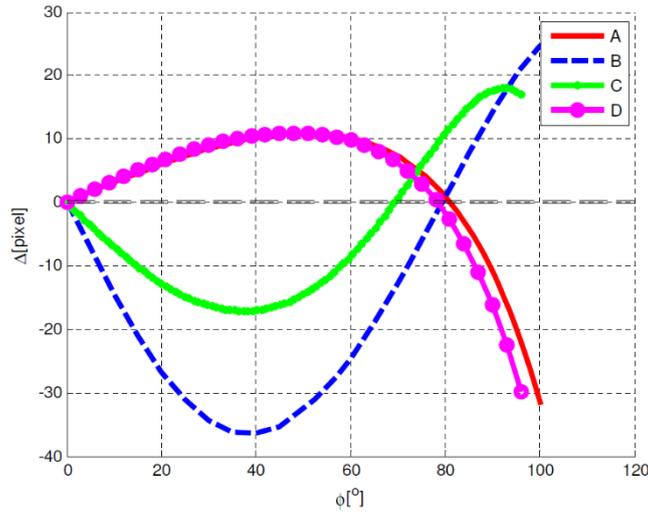


Fig. 13.2. The results of equidistance model (Index 1) after calibration, Deviation from design data

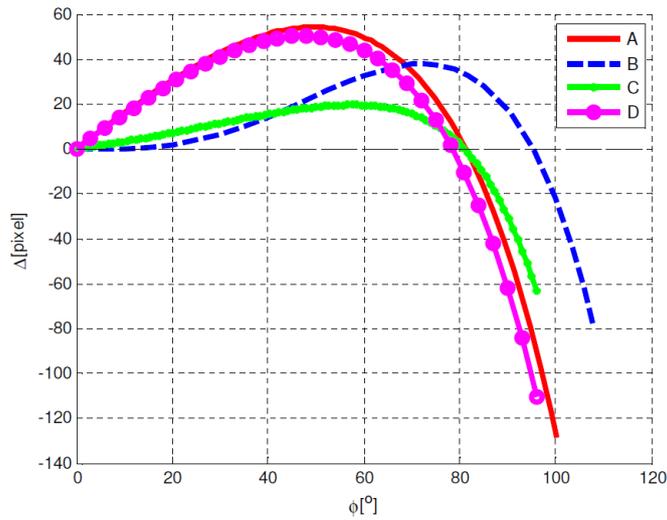


Fig. 13.3. The results of stereographic model (Index 2) after calibration, Deviation from design data

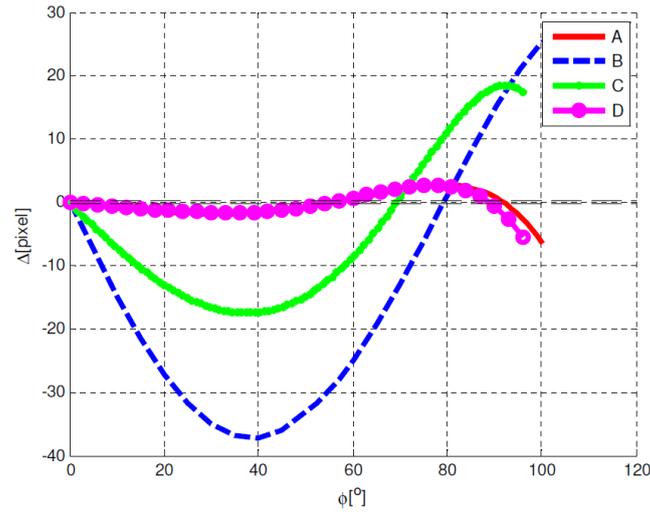


Fig. 13.4. The results of modified orthogonal model (Index 3) after calibration, Deviation from design data

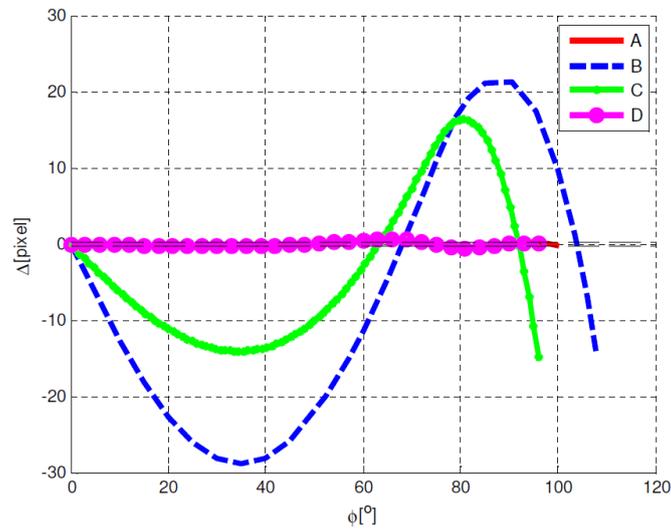


Fig. 13.5. The results of modified orthogonal extended model (Index 4) after calibration, Deviation from design data

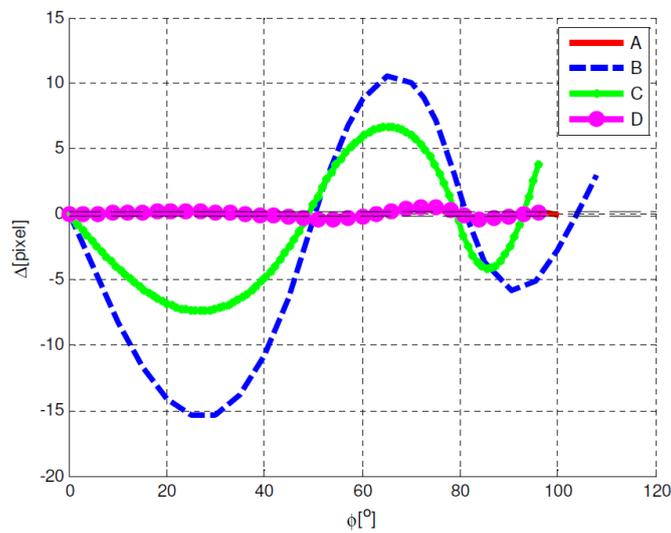


Fig. 13.6. The results of modified orthogonal extended model (Index 5) after calibration, Deviation from design data

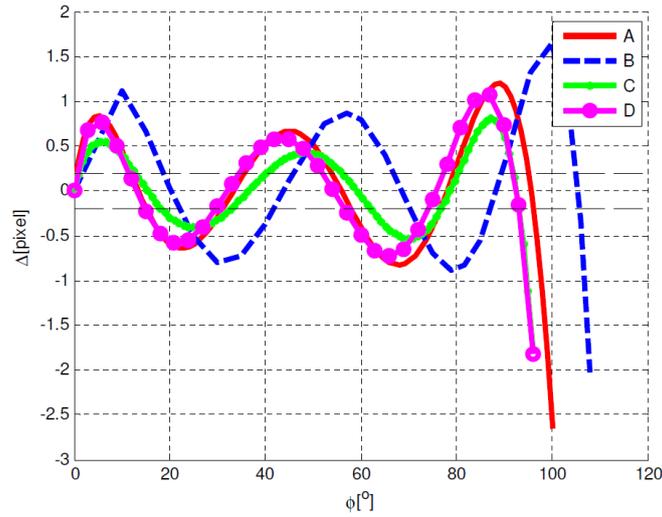


Fig. 13.7. The results of Scaramuzza model (Index 6) after calibration, Deviation from design data

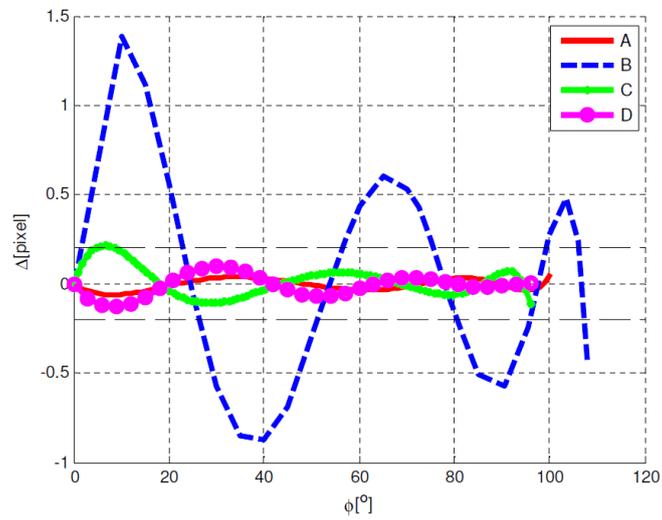


Fig. 13.8. The results of polynomial model (Index 7) after calibration, Deviation from design data

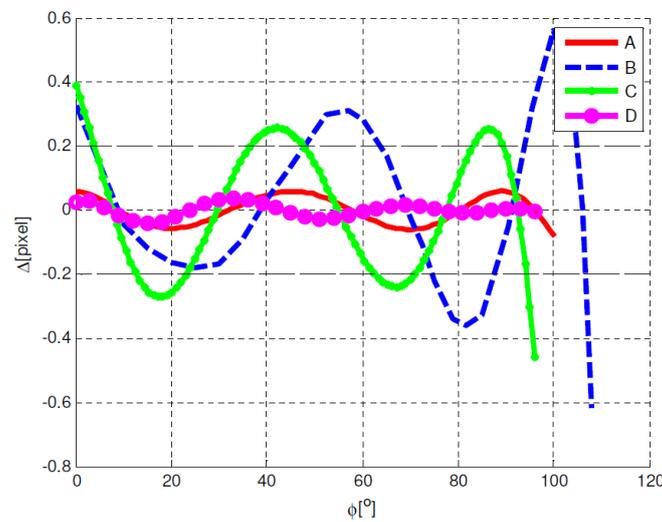


Fig. 13.9. The results of Tschebycheff model (Index 8) after calibration, Deviation from design data

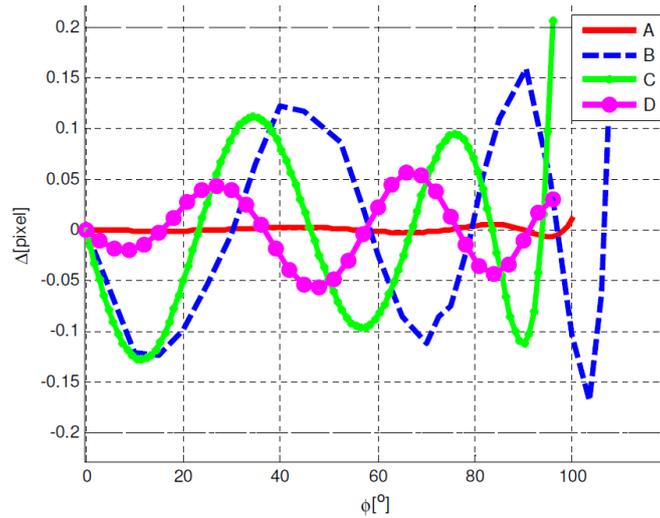


Fig. 13.10. The results of Christopher Mei model (Index 9) after calibration, Deviation from design data

Now we can compare the maximum error in all of the models. The results are illustrated in Fig. 13.11:

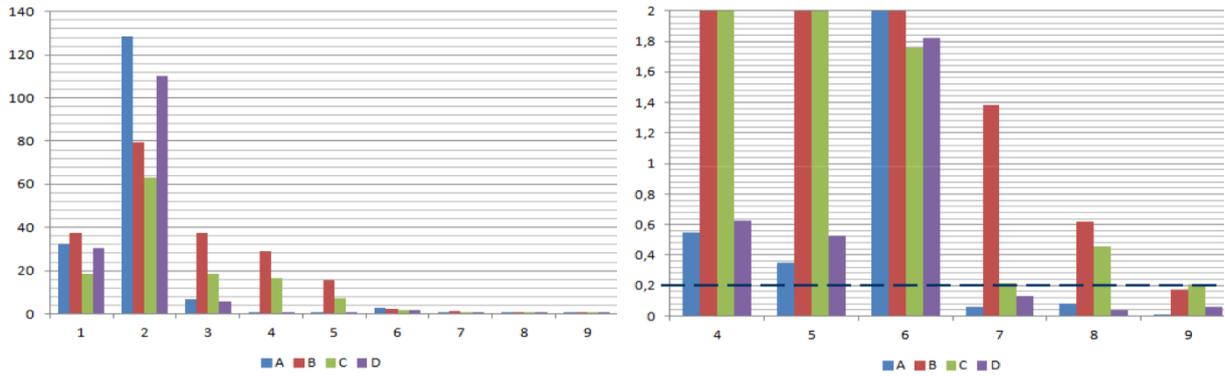


Fig. 13.11. The results of all models (index 1..9) after calibration, maximum deviation from design data

As we can see, Christopher Mei model (Index 9) can provide a decent accuracy with the maximum error smaller than 0.2 [Pixel] using the design data for all the cameras. Therefore, this model is considered as a very suitable model for accurately calibrating fisheye cameras.

14. Appendix: Abbreviations

AP	additional parameters
C2W	camera to world
CDF	cumulative distribution function
CPU	central processing unit
DA	driver assistant (system)
DIN	Deutsches Institut für Normung (German Standard Institute)
DLT	direct linear transformation
DOF	degree of Freedom
FOV	field of view
FPGA	field-programmable gate array
fps	frames per second
GPU	graphic processing unit
GPGPU	general purpose GPU
GPS	global positioning system
GSD	ground sampling distance, sampling distance in object space
ISP	image signal processing
LED	light emitting diode
LoG	Laplacian of Gaussian
LSE	least squares estimation
LUT	lookup table
MPC	multi-purpose camera
MTF	modulated transfer function
NRC	near-range camera
PDF	probability distribution function
RGB	red, green, blue
RMS	root mean square
RMSE	root mean square of errors
SFM	structure from motion
SRVC	smart reverse view camera
SVC	stereo-video camera
TDP	thermal design power
W2C	world to camera
3DTV	three dimensional top-view

15. Appendix: Conventions

§	paragraph
@	at, in case of, with the value of
CTE, cte	constant
Fig.	Figure
Eq.	Equation
∀	Each
&	logical AND
	logical OR
..	to (for domain definition)
∝	is proportional to
≈	approximately equal
ε	epsilon, very small number, almost ignorable, remaining error
.	representing multiplications
[mm]	unit description (e.g. [mm]: millimetres)
[1.2.1]	reference to chapter (e.g. [1.2.1]: chapter number 1.2.1)
[A..B]	domain definition (from A to B)
{A..B}	domain definition (from A to B)
<< or >>	intensified comparative
:=	value setting, calculation of left part from the right part
...	etc, and so on
\hat{A}	posterior/estimated value for A
∂	used for partial differentiation
d	used for derivations, small increments/decrements
Δ	increments/decrements
A^t	transpose operator for matrix A
A^{-1}	inverse of matrix A
f^{-1}	inverse of function f
⇒	Therefore / as a result / thereupon

Bibliography

- Abraham, S., 2004. *A Software for Test-Field Based Self-Calibration of Multi-Camera Systems*, s.l.: Institute for Photogrammetry, University of Bonn.
- Abraham, S., 2012. *Evaluierung mathematischer Modelle zur Beschreibung der Fish-Eye Verzeichnung für NRC, Untersuchung an Hand des optischen Designs*, Hildesheim, Germany: Robert Bosch GmbH, Chassis Control Systems, Engineering System and Software - Video (ESV), (Ref. available in Appendix).
- Abraham, S. & Förstner, W., 2005. *Fish-Eye-Stereo Calibration and Epipolar Rectification*, s.l.: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 59, Issue 5, pp 278–288, doi: 10.1016/j.isprsjprs.2005.03.001.
- Abraham, S. & Hau, T., 1997. *Toward Autonomous High-Precision Calibration of Digital Cameras*, Bonn University: SPIE 3174 Annual Meeting, Videometrics V, doi:10.1117/12.279802.
- Ahn, S. J. & Kotowski, R., 1997. *Geometric Image Measurement Errors of Circular Object Targets*, ETH Zurich, Switzerland: 4th Conference on Optical 3-D Measurement Techniques.
- Amini, J., 2006. *Analytical Photogrammetry, (in Persian)*. 2nd Edition ed. s.l.:University of Tehran.
- Amiri-Seemkooei, A., 2001. *Strategy for Designing Geodetic Network with High Reliability and Geometrical Strength*, s.l.: Journal of Surveying Engineering (JSE), 127, pp 104-117, DOI: 10.1061/(ASCE)0733-9453(2001)127:3(104).
- Amiri-Simkooei, A., Asgari, J., Zangeneh-Nejad, F. & Zaminpardaz, S., 2012. *Basic Concepts of Optimization and Design of Geodetic Networks*, s.l.: Journal of Surveying Engineering (JSE). 138. pp 172-183, DOI:10.1061/(ASCE)SU.1943-5428.0000081.
- Apostolico, A., Browne, S. & Guerra, C., 1992. *Fast linear-space computation of longest common subsequences*, s.l.: Theoretical Computer Science 92, Elsevier, pp 3-17, doi: 10.1016/0304-3975(92)90132-Y.
- Baretto, J. P. & Daniilidis, K., 2004. *Wide Area Multiple Camera Calibration and Estimation of Radial Distortion*, Prague: Omnivis2004, The fifth Workshop on Omnidirectional Vision, Camera Networks and Non-classical cameras.
- Bass, M., Van Stryland, E. W., Williams, D. R. & Wolfe, W. I., 1995. *Handbook of optics Fundamentals, Techniques, & Design*. 2nd Edition ed. s.l.:McGraw-Hill, Inc, ISBN 007047740X, 9780070477407.
- Benhenni, k. & Degras, D., 2011. *Local Polynomial Regression Based on Functional Data*, s.l.: Journal of Multivariate Analysis, arXiv: 1107.4058.
- Born, M. & Wolf, E., 1999. *Principles of Optics*. 7th (expanded) edition ed. Cambridge: Cambridge University Press, ISBN 10: 0521642221, ISBN 13: 9780521642224.

- Brown, D., 1976. *The bundle method - progress and prospects*, Helsinki: International Archives of Photogrammetry, Volume 21(3), pp 1-33, ISP congerss.
- Brown, D. C., 1971. *Close-Range Camera Calibration*, s.l.: Photogrammetric Engineering, Volume 37.
- Cefalu, A. & Böhm, J., 2010. *Development of a Robot Guided Optical Multisensory Inspection System for Inline Inspection of Cylinder Heads*, Newcastle upon Tyne, UK: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII, Part 5 Commission V Symposium, pp 137-142.
- Censi, A. & Scaramuzza, D., 2013. *Calibration by correlation using metric embedding from non-metric similarities*, s.l.: IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol 35, 2357-70. 10.1109/TPAMI.2013.34..
- Clarke, T. A. & Fryer, J. G., 1998. *The development of camera calibration methods and models*, s.l.: The Photogrammetric Record, Volume 16, Issue 91, pp 51-66.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C., 2009. *Introduction to algorithms*. Third edition ed. s.l.:The MIT Press, ISBN-10: 0262533057, ISBN-13: 978-0262533058.
- Cross, P. A., 1985. *Numerical Methods in Network Design*. In: Grafarend E.W., Sansò F. (eds) *Optimization and Design of Geodetic Networks*, Berlin, Heidelberg: Springer, DOI: 10.1007/978-3-642-70659-2_7.
- Dahlquist, G. & Björck, Å., 2003. *Numerical Methods*. s.l.:Dover Publication, Inc. (Prentice-Hall, Inc. 1974), pp 101-103.
- De Villiers, J. P., Leuschner, F. W. & Geldenhuys, R., 2008. *Centi-pixel accurate real-time inverse distortion correction*, s.l.: Proceedings of SPIE Vol. 7226, doi:10.1117/12.804771.
- Ebner, H., 1976. *Self-calibration block adjustment*, Helsinki: XIII Congress of the International Society for Photogrammetry, International Archives of Photogrammetry, XXI/3.
- Estribeau, M. & Mignan, P., 2004. *Fast MTF measurement of CMOS imagers using ISO 12333 slanted-edge methodology*, s.l.: Proc. SPIE 5251, Detectors and Associated Signal Processing, 243, doi:10.1117/12.513320.
- EUR-Lex, I. 4.-1., 1978. *ISO 4130-1978, RICHTLINIE DES RATES (vom 27. September 1977), zur Angleichung der Rechtsvorschriften der Mitgliedstaaten über das Sichtfeld der Fahrer von Kraftfahrzeugen (77/649/EWG)*, s.l.: EUR-Lex.
- Fraser, C. S., 1997. *Digital Camera Self-Calibration*, s.l.: ISPRS Journal of Photogrammetry and Remote Sensing, Volume 52, Issue 4, pp 149-159, ISSN 0924-2716, doi:10.1016/S0924-2716(97)00005-1.
- Fritzpatrick, R., 2007. *Electromagnetism and Optics, an introductory course*, s.l.: University of Texas, Institute for Fusion Studies.
- Fryer, J. G. & Brown, D. C., 1986. *Lens distortion for close-range photogrammetry*, s.l.: Photogrammetric Engineering and Remote Sensing, 52, pp 51-58..

- Fuqiang, Z., Yi, C., He, G. & Yexin, W., 2013. *Line-based camera calibration with lens distortion correction from a single image*, s.l.: Optics and Lasers in Engineering, Volume 51, Issue 12, Pages. 1332-1343, ISSN 0143-8166, doi:10.1016/j.optlaseng.2013.05.010.
- Gonzalez, R. C. & Woods, R. E., 2007. *Digital Image Processing*. 3rd Edition ed. s.l.:Prentice Hall International, ISBN-10: 013168728X, ISBN-13: 978-0131687288.
- Grafarend, E. W., 1974. *Optimization of geodetic networks*, s.l.: Bolletino di Geodesiaa Science Affini, 33(4), 351–406.
- Grafarend, E. W. & Sanso, F., 1985. *Optimization and design of geodetic networks*. Berlin: Springer.
- Grün, A., 1978. *Progress in photogrammetric point determination by compensation of systematic errors and detection of gross errors*, Moscow: Symposium of Comm. III of the ISP (113-140).
- Guest, P. G., 1958. *The Spacing of Observations in Polynomial Regression*, University of Sydney, Australia: The Annuals of Mathematical Statistics, Vol. 29, No. 1, 1958, pp. 294–299. JSTOR.
- Haala, N., 2011. *Multiray Photogrammetry and Dense Image Matching*, s.l.: Photogrammetric Week, Wichmann Verlag, Berlin/Offenbach, pp 185-195.
- Hartley, R. & Zisserman, A., 2011. *Multiple View Geometry in Computer Vision*. 2nd Edition ed. s.l.:Cambridge University Press, ISBN: 9780511811685.
- Hernández, C. & Furukawa, Y., 2013. *Multi-View Stereo: A Tutorial*. Vol. 9, No. 1-2 (2013) 1–148 ed. s.l.:Now Publishers, Inc, ISBN: 9781601988362.
- Hoschek, J. & Lasser, D., 1992. *Scattered Data Interpolation*. In: *Grundlagen der geometrischen Datenverarbeitung*. s.l.:Vieweg+Teubner Verlag, ISBN: 978-3-322-89829-6.
- Hosoya, H., 2011. *High π -Electronic Stability of Soccer Ball Fullerene C60 and Truncated C24 Among Spherically Polyhedral Networks*. s.l.:The Mathematics and Topology of Fullerenes, Springer, Chapter 13, pp 249-263, 10.1007/978-94-007-0221-9_13.
- Hummel, K. E., 2000. *Introductory Concepts for Abstract Mathematics*. s.l.:Chapman & Hall/CRC, pp 195-197, ISBN-10: 1584881348, ISBN-13: 978-1584881346.
- Kirkwood, J. R., 2002. *An Introduction to Analysis*. 2nd Edition ed. s.l.:Waveland Press, ISBN-10: 1577662326, ISBN-13: 978-1577662327.
- Kreyszig, E., 2010. *Advanced Engineering Mathematics*. 10th Edition ed. s.l.:John Wiley & Sons Ltd, ISBN-10: 0470458364, ISBN-13: 978-0470458365.
- Krüger, T. & Meinel, G., 2008. *in: Advances in 3D Geoinformation Systems*. s.l.:Springer, Berlin, Heidelberg, ISBN: 978-3-540-72135-2, doi: 10.1007/978-3-540-72135-2_6.
- Kuang, S. L., 1996. *Geodetic network analysis and optimal design: Concepts and application*. Chelsea, MI.: Ann Arbor Press.

- Laderman, J. & Laderman, J. D., 1982. *Simplified Forecasting by Polynomial Regression with Equally Spaced Values of the Independent Variable*, s.l.: American Mathematical Society, Mathematics of Computation, Volume 38, Number 158, pp. 601-610, doi: 10.2307/2007294.
- Lichti, D. D. & Chapman, M. A., 1995. *CCD camera calibration using the finite element method*, s.l.: Proc. SPIE, Vol 2598, Videometrics IV, pp 33-43, doi: 10.1117/12.220911.
- Luhmann, T., 2010. *Erweiterte Verfahren zur geometrischen Kamerakalibrierung in der Nahbereichsphotogrammetrie, Dissertation*. München: Verlag der Bayerischen Akademie der Wissenschaften in Kommission beim Verlag C. H. Beck, ISSN 0065-5325, ISBN 978-3-7696-5057-0.
- Luhmann, T., Bobson, S., Kyle, S. & Harley, I., 2011. *Close Range Photogrammetry, Principles, Methods and Applications*. s.l.:Whittles Publishing, ISBN-10: 1849950571, ISBN-13: 978-1849950572.
- Mahajan, V. N., 1998. *Optical Imaging and Aberrations, Part 1, Ray Geometrical Optics*. The Aerospace Corporation and the University of Southern California: SPIE Press, ISBN 9780819425157.
- Mallon, J. & Whelan, P. F., 2004. *Precise Radial Un-Distortion of Images*, s.l.: Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), IEEE Computer Society, doi: 10.1109/ICPR.2004.1333995.
- Mason, J. C. & Handscomb, D. C., 2002. *Chebyshev Polynomials*. s.l.:Chapman & Hall / CRC, ISBN 9780849303555.
- Mazaheri, M. & Momeni, M., 2008. *3D Modeling Using Structured Light Pattern and Photogrammetric Epipolar Geometry*, s.l.: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 37.
- Mei, C. & Rives, P., 2007. *Single View Point Omnidirectional Camera Calibration from Planar Grids*, s.l.: Proceedings 2007 IEEE International Conference on Robotics and Automation, pp 3945-3950, doi: 10.1109/ROBOT.2007.364084.
- Moffitt, F. H. & Mikhail, E. M., 1980. *Photogrammetry*. 3rd Edition ed. s.l.:Harper & Row, Publishers, New York, ISBN: 070022517X, 9780700225170.
- Moussa, W. & Fritsch, D., 2010. *A Simple Approach to Link 3D Photorealistic Models with Content of Bibliographic Repositories*, s.l.: Springer, Berlin, Heidelberg, Vol. 6436 of the series Lecture Notes in Computer Science, pp 482-491, ISBN: 978-3-642-16873-4, doi: 10.1007/978-3-642-16873-4_39.
- Nekouei Shahraki, M., 2012. *Camera Alignment Testing and Calibration, Master Thesis Report*, Stuttgart, Germany: Robert Bosch GmbH, ifp - University of Stuttgart.
- Nekouei Shahraki, M. & Haala, N., 2015 A. *Introducing free-function camera caibration model for central-projection and omni-directional lenses*, s.l.: Proc. SPIE 9630, Optical Systems Design, doi: 10.1117/12.2191121.
- Nekouei Shahraki, M. & Haala, N., 2015 B. *Introducing Novel Generation of High Accuracy Camera Optical-Testing and Calibration Test-Stands Feasible for Series Production of Cameras*, s.l.:

Copernicus, ISPRS, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Volume XL-1-W5, pp 521-527, doi: 10.5194/isprsarchives-XL-1-W5-521-2015.

Niederöst, M. & Maas, H.-G., 1997. *Entwurf und Erkennung von codierten Zielmarken*, Oldenburg, Germany: Publikation der DGPF, Band 5, 16. Wissenschaftlich-Technische Jahrestagung der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung (DGPF), doi: 10.3929/ethz-a-004332841.

Parasad, L. & Lyengar, S. S., 1997. *Wavelet Analysis with Applications to Image Processing*. s.l.:CRC Press LLC, ISBN-10: 0849331692, ISBN-13: 978-0849331695.

Payne, C., Bjore, C., Cromey, D. & Roland, F., 1989. *A Comparative Mathematical Evaluation of Contour Irregularity Using Form Factor and PERBAS, A New Analytical Shape Factor*, s.l.: Analytical and Quantitative Cytology and Histology, Vol. 11(5), pp 341-352. PMID: 2679617.

Pottmann, H., Hofer, M., Odehnal, B. & Wallner, J., 2004. *Line Geometry for 3D Shape Understanding and Reconstruction, Classification of Surfaces by Normal Congruences*, Prague, Czech Republic: Springer, Computer Vision – ECCV: 8th Europea Conference on Computer Vision, doi: 10.1007/978-3-540-24670-1_23.

Pottmann, H., Wallner, J. & Leopolseder, S., 2002. *Kinematical methods for the classification, reconstruction and inspection of surfaces*, Corr`eze, France: Comptes rendus du Congr`es national de math´ematiques appliqu´ees et industrielles, pp 51-60.

Prescott, B. & McLean, G. F., 1997. *Line-Based Correction of Radial Lens Distortion*, s.l.: Graphical Models and Image Processing, Vol. 59, Issue 1, Pages 39-47, ISSN 1077-3169, doi: 10.1006/gmip.1996.0407.

Raizner, C., 2012. *Objective and Automated Stray Light Inspection of High-Dynamic-Range Cameras*, s.l.: Shaker Verlag, Herzogenrath, ISBN 978-3-8440-1278-3, 240 S.

Ranganathan, P. & Olson, E., 2012. *Gaussian Process for Lens Distortion Modeling*, Vilamoura, 2012: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 3620-3625, doi: 10.1109/IROS.2012.6385481.

Reznicek, J. & Luhmann, T., 2019. *Finite-Element Approach to Camera Modelling and Calibration*, s.l.: Springer, PFG 87, pp 1-17, doi: 10.1007/s41064-019-00068-x.

Reznicek, J., Luhmann, T. & Jepping, C., 2016. *Influence of raw image preprocessing and other selected processes on accuracy of close-range photogrammetric systems according to VDI 2634*, s.l.: ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XLI-B5, pp 107-113, doi: 10.5194/isprsarchives-XLI-B5-107-2016.

Richter, K., Mader, D., Seidl, K. & Maas, H.-G., 2013. *Development of a geometric model for an all-reflective camera system*, s.l.: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 86, pp 41-51, ISSN 0924-2716, doi: 10.1016/j.isprsjprs.2013.09.002..

Scaramuzza, D., 2014. *Omnidirectional Camera*, s.l.: Computer Vision, Springer, pp 552-560, doi: 10.1007/978-0-387-31439-6_488.

- Scaramuzza, D., Martinelli, A. & Siegwart, R., 2006. *A Tool-Box for Easily Calibrating Omnidirectional Camera*, s.l.: Intelligent Robots and Systems, IEEE/RSJ International Conference, doi: 10.1109/IROS.2006.282372.
- Scheuß, U., 1995. *Entwicklung eines autonomen Kalibrierungsverfahrens für das Stereokamerasystem eines Roboters*, Master Thesis, s.l.: University of Bonn.
- Schneider, C.-T., 1996. *A PC Based Digital Photogrammetric Station for Fast and Flexible Onsite Measurement*, Vienna: International Archives of Photogrammetry and Remote Sensing, Vol. XXI, Part B5, pp 530–533.
- Schneider, D., Schwalbe, E. & H.-G., M., 2009. *Validation of geometric models for fisheye lenses*, s.l.: ISPRS Journal of Photogrammetry and Remote Sensing, Vol 64, Issue 3, pp 259-266, ISSN 0924-2716, doi: 10.1016/j.isprsjprs.2009.01.001..
- Schönberger, J. L. & Frahm, J.-M., 2016. *Structure-from-Motion Revisited*, Las Vegas, USA: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4104-4113, ISSN 1063-6919, doi: 10.1109/CVPR.2016.445.
- Schröder, G., 1981. *Technische Fotografie: Grundlagen und Anwendungen in Technik und Wissenschaft*. s.l.:Vogel-Verlag, ISBN-10: 3834332186, ISBN-13: 978-3834332189.
- Schröder, G. & Treiber, H., 2002. *Technische Optik, Grundlagen und Anwendungen (in German)*. Edition 9, erweiterte Auflage ed. s.l.:Vogel Verlag, ISBN: 3802319230, 9783802319235.
- Sebah, P. & Gourdon, X., 2011. *Newton's Method and High Order Iterations*, s.l.: Institute for Computer Science and Control, Hungarian Academy of Sciences (SZTAKI), <http://old.sztaki.hu/~bozoki/oktatas/nemlinearis/SebahGourdon-Newton.pdf>.
- Shephard, D., 1968. *A Two-Dimensional Interpolation Function For Irregularly-Spaced Data*, s.l.: ACM 68, Proceedings of the 1968 23rd ACM national conference, pp 517-527, doi: 10.1145/800186.810616.
- Stakhov, A., 2009. *The Mathematics of Harmony: From Euclid to Contemporary Mathematics and Computer Science*. s.l.:World Scientific Publishing Company, Vol. 22, pp 144-146, ISBN-10: 981277582X, ISBN-13: 978-9812775825.
- Szeliski, R., 2010. *Computer Vision: Algorithms and Applications*. s.l.:Springer, ISBN-10: 1848829345, ISBN-13: 978-1848829343.
- Tang, R., Fritsch, D. & Cramer, M., 2012 A. *A Novel Family of Mathematical Self-Calibration Additional Parameters for Airborne Camera Systems*, s.l.: European Calibration and Orientation Workshop (EuroCOW), 7 pages on CD-ROM.
- Tang, R., Fritsch, D. & Cramer, M., 2012 B. *New Rigorous and Flexible Fourier Self-Calibration Models for Airborne Camera Calibration*, ifp - University of Stuttgart: ISPRS Journal of Photogrammetry and Remote Sensing, Vol 71, pp 76-85, doi: 10.1016/j.isprsjprs.2012.05.004.

- Weber, S., 2012. *Verfahren und Vorrichtung zur Bestimmung von zur Entzerrung eines Bildes geeigneten Werten*, Germany: Deutsche Patent- und Markenamt (Patent), Nr. DE102011007644A1.
- Westboy, M. J. et al., 2012. '*Structure-from-Motion*' *Photogrammetry: A Low-Cost, Effective Tool for Geoscience Applications*, s.l.: Elsevier, *Geomorphology*, Vol. 179, pp 300-314, ISSN 0169-555X, doi: 10.1016/j.geomorph.2012.08.021.
- Xiong, Y. & Turkowski, K., 1997. *Creating Image-Based VR Using a Self-Calibrating Fisheye Lens*, San Juan, Puerto Rico, USA: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 237-243, doi: 10.1109/CVPR.1997.609326.
- Ying, X., Hu, Z. & Zha, H., 2006. *Fisheye Lenses Calibration Using Straight Line Spherical Perspective Projection Constraint*, s.l.: In: Narayanan P.J., Nayar S.K., Shum HY. (eds) *Computer Vision – ACCV 2006*. Lecture Notes in Computer Science, Vol. 3852, Springer, Berlin, Heidelberg, ISBN: 978-3-540-32432-4, doi: 10.1007/11612704_7.

Curriculum Vitae

Mehrdad, Nekouei Shahraki

Address: Institut für Photogrammetrie, Geschwister-Scholl-Str. 24D, 70174, Stuttgart

Birthday: 21 Sep 1987

Birthplace: Shahrekord, Iran

Tel: +49-176-30116002

Email: Mehrdad.NekoueiShahraki@de.bosch.com

Mehrdad.Nekouei@gmail.com

EDUCATION

2005-2009	Bachelor	Faculty of Engineering, Department of Geomatics, University of Isfahan
2010-2012	Master	GeoEngine Program, University of Stuttgart
2012-	Ph.D.	Institute for Photogrammetry, University of Stuttgart

EXPERIENCE

2007-2009	Research Assistant	University of Isfahan
2011-2012	Master Student	Robert Bosch GmbH
2012-2015	Doctorand (Ph.D. Associate)	Robert Bosch GmbH
2015-	Engineering Associate	Robert Bosch GmbH