

MONOCULAR CAMERA CALIBRATION ASSESSMENT FOR MID-RANGE PHOTOGRAMMETRY

Jean-Sébastien Valois

Department of Electrical Engineering
McGill University

June 2000

A Thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the degree of
Master of Engineering

© JEAN-SÉBASTIEN VALOIS, MM

Abstract

The CCD (Charge Coupled Device) image formation theory is at the foundation of 3D vision systems. Ideally, a five-element block diagram can model this process. The first block represents the nonlinear distortions caused by camera lenses. The second and third elements gather the low-pass filter effects due to lens aberrations and CCD phenomenon. A fourth block illustrates the quantization effects induced by a series of discrete photosensitive elements on the CCD and by the A/D conversion for analog cameras. The last block represents the addition of random noise on the discrete signal. Because step-like luminance transitions undergoing lens distortion remain step-like, it is possible to precisely correct for the distortions after the edge localization. The efficient distortion correction process is exactly where the camera lens calibration challenge resides. The calibration exercise also seeks the intrinsic and extrinsic camera parameters, i.e. the information that relates to the camera optics and the information that describes the location and orientation of the camera in 3D space.

This thesis presents a review and evaluation of several methods designed for optimal accuracy on the parameters evaluation. We also discuss the aspects affecting the accurate recovery of the camera perspective geometry parameters which are affected by numerous error sources.

The techniques studied can be classified into two distinct calibration categories: the *generic* and the *parametric* methods. The parametric calibrations consist of determining and solving a set of nonlinear mathematical equations dependent on the knowledge of the physical interaction between a camera, a lens and features in the 3D world. The main advantage is that no approximation is involved and the model can be quite elaborate. The problem results in the complexity to perform the nonlinear optimization and the need for

a good initial guess. With the generic calibration methods the explicit expression of the distortions is not assumed. Also, they do not necessarily aim at computing all the camera parameters simultaneously since once the distortions are handled, linear approaches can perform this task very well.

The performance of the different calibrations was evaluated at two levels. The first level verifies the validity and accuracy of the distortion corrections using the calibration equipment itself. The second one follows a field test approach where a volumetric array of size similar to the real operation environment is used to verify the accuracy of both the intrinsic and extrinsic camera parameters.

The thesis also includes a thorough discussion on the equipment used for calibrating cameras. The quality of the calibration equipment has a critical impact on the camera calibration accuracy.

Résumé

La théorie de formation d'images sur CCD (Charge Coupled Device) est à la base des systèmes de vision tridimensionnelle. Idéalement, un diagramme composé de cinq éléments peut représenter le phénomène complètement. Le premier bloc représente les distorsions de type non linéaire causées par les lentilles. Le second et troisième blocs regroupent les effets de filtre à haute fréquence qui sont due à la fois aux lentilles et au capteur CCD. Un quatrième bloc illustre l'effet d'approximation introduit par l'échantillonnage de la lumière par les éléments photo sensitifs du CCD et par la conversion A/D. Le dernier bloc représente l'addition de bruit sur le signal discret. Le fait que les transitions vives d'intensité dans l'image ne soient pas affectées par la distorsion des lentilles permet, après avoir localisé ces transitions, de corriger ces distorsions. Le défi réside donc dans la correction efficace des distorsions causées par les lentilles de caméras. Par le fait même, l'exercice de calibrage de caméras consiste aussi à chercher les paramètres intrinsèques et extrinsèques de la camera, c'est-à-dire, l'information sur l'aspect optique interne de la camera et l'information décrivant sa location et son orientation dans un environnement tridimensionnel.

Cette thèse présente une revue et une évaluation de plusieurs méthodes qui furent introduites pour maximiser la précision des paramètres lors du calibrage. Nous abordons également les phénomènes qui affectent la reconstitution des paramètres décrivant la géométrie perspective de la caméra.

Les techniques étudiées peuvent être regroupées en deux catégories: les méthodes *génériques* et les méthodes *paramétriques*. Les méthodes paramétriques consistent à déterminer et résoudre un ensemble d'équations non linéaires qui dépendent de connaissances sur les aspects physiques d'interaction entre les lentilles, la caméra et l'environnement tridimensionnel. Le principal avantage de ces méthodes est qu'aucune approximation est nécessaire et que le

modèle peut être très élaboré. Par contre, certains problèmes proviennent de la complexité nécessaire pour résoudre les équations et le besoin d'avoir un point de départ proche de la solution finale. Pour ce qui est des méthodes dites génériques, l'expression explicite des distorsions n'est pas nécessaire. De plus, ces méthodes ne cherchent pas nécessairement à retrouver tous les paramètres extrinsèques et intrinsèques d'un seul coup. Plutôt, les distorsions optiques sont d'abord corrigées et par la suite le reste des paramètres de la caméra peut être facilement obtenu.

La performance des diverses méthodes peut être évaluée à deux niveaux. Premièrement, le niveau de précision de correction des distorsions est vérifié en utilisant l'équipement de calibrage lui-même. En second lieu, une approche s'apparentant aux conditions réelles d'opération est utilisée. Ainsi, un arrangement de cibles de tailles et à une distance similaire aux opérations sert à mesurer l'ensemble des paramètres extrinsèques et intrinsèques de la caméra.

Cette thèse contient également une discussion complète sur l'équipement et les conditions nécessaires pour réaliser de bons calibrages.

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my advisor, Professor Frank P. Ferrie, for his help, support and understanding throughout my Master's degree.

I would also like to state my deepest admiration and gratitude to Steve G. MacLean for believing in me and for introducing me to the field of Computer Vision. His guidance has changed the course of my career.

I would also like to acknowledge and thank my colleagues at Neptec Design Group and NASA. In particular, Richard Theobald, Larisa Beach, Erik Lockeberg, Lloyd Pinkney, Mike Sink, Larry Schmidt, David Moyer and Jim Newman. Their enthusiasm provides such an amazingly fertile environment to work in.

I would also like to thank my parents for their unwaivering encouragement and support. And finally to my wife Kerry, whose moral support I am very grateful for.

TABLE OF CONTENTS

Abstract	i
Résumé	iii
ACKNOWLEDGMENTS	v
LIST OF FIGURES	xi
LIST OF TABLES	xiv
CHAPTER 1. Introduction	1
1. Background	1
2. Problem definition	5
2.1. The camera calibration challenges	5
3. The state of the art	7
3.1. The parametric calibration techniques	8
3.2. The generic calibration techniques	14
3.3. Different calibration methods, same goal	17
3.4. Calibration accuracy evaluation techniques	17
3.5. The most common calibration objects	18
4. Main contributions	18
5. Thesis outline	20
CHAPTER 2. General image formation concept	22
1. Lens system aberrations	23
1.1. Geometrical distortions (monochromatic effect)	25

1.2.	Spherical aberrations (monochromatic effect)	26
1.3.	Coma (monochromatic effect)	26
1.4.	Field curvature (monochromatic effect)	27
1.5.	Astigmatism (monochromatic effect)	27
1.6.	Longitudinal/Axial Color (chromatic effect)	30
1.7.	Lateral Color (chromatic effect)	30
1.8.	General lens aberration contributions	31
2.	CCD sensor aberrations	33
2.1.	CCD operation	33
2.2.	Video frame grabbing	35
2.3.	Spatial resolution	35
2.4.	Dynamic resolution	36
2.5.	Sensitivity	36
2.6.	Photon noise	36
2.7.	Dark current and dark current noise	37
2.8.	CCD defects	37
2.9.	Clocking noise	37
2.10.	Reset noise	38
2.11.	Charge transfer efficiency	38
2.12.	CCD readout and charge conversion limitations	38
2.13.	Charge-to-voltage conversion limitations	38
2.14.	Video encoding	39
2.15.	General CCD aberration contributions	39
3.	CCD and lens contributions	39
4.	Conclusion	40
CHAPTER 3.	Mathematical concepts for camera calibration	41
1.	The camera model and coordinate systems for the calibration	41
1.1.	Rigid body transformation from the TCS to the CCS	43
1.2.	Transformation from the CCS to ideal undistorted ICS	46
1.3.	Lens distortion correction	46
1.4.	Transformation from the FCS to the ICS	47

2. Extrinsic and intrinsic camera parameters	49
2.1. The extrinsic camera parameters	50
2.2. The intrinsic camera parameters	50
3. Conclusion	51
CHAPTER 4. Camera calibration requirements	52
1. Camera calibration equipment	52
1.1. Construction tolerances	53
1.2. Calibration array pattern	55
1.3. 2D versus 3D calibration objects	61
2. Camera configuration	65
3. Laboratory conditions	68
4. Data recording	69
5. Calibration setup used for this thesis	70
5.1. Camera description	70
5.2. Calibration rig	71
5.3. Data collection and processing	72
6. Conclusion	73
CHAPTER 5. Camera calibration methods tested	74
1. Generic calibration methods	74
1.1. Original method	75
1.2. Lookup compensation method	82
1.3. Averaging of multiple compensated maps	84
1.4. Polynomial smoothing method	86
1.5. Polynomial smoothing with lookup compensation method	87
2. Parametric calibration methods	87
2.1. The solution process for the parametric methods	88
2.2. Parametric distortion models	89
2.3. Parametric distortion models tested	92
3. Conclusion	96
CHAPTER 6. Calibration performance evaluation	97

1. Estimated centroid error	97
2. Calibration evaluation on the calibration board	98
2.1. Camera-Space Error (<i>CSE</i>) on the calibration board	100
3. Calibration evaluation on the volumetric array	103
3.1. Camera-Space Error (<i>CSE</i>) on the volumetric array	104
3.2. Object-Space Error (<i>OSE</i>) on the volumetric array	106
3.3. Relative photosolution error (<i>RPSE</i>) on the volumetric array	107
4. Conclusion	109
CHAPTER 7. Test results	111
1. Results of the <i>CSE</i> tests on the calibration board	112
2. Results of the <i>CSE</i> tests on the volumetric array	113
3. Results of the <i>OSE</i> tests on the volumetric array	114
4. Results of the <i>RPSE</i> tests on the volumetric array	114
5. Performance of the different calibration methods	115
5.1. Generic methods review	116
5.2. Parametric methods review	118
6. General use of the calibration evaluation methods	120
7. Conclusion	120
CHAPTER 8. Conclusion	121
1. Preferred calibration method	122
2. Camera configuration	122
3. Calibration rig	122
4. Camera calibration evaluation	123
5. Areas of future research	123
GLOSSARY	125
REFERENCES	127
APPENDIX A. Camera configurations	133
1. CTVC	133
2. ITVC	134

TABLE OF CONTENTS

3. Cohu 4910	136
APPENDIX B. Least-squares minimization method	138
1. <i>Radial - first order</i> example	138
APPENDIX C. Student's t-test results	143
1. Student's t-test in a nutshell	143
2. Complete calibration performance results	143

LIST OF FIGURES

1.1	Scene from the first ISS assembly mission (STS-88). The Zarya module (Russia) had just been joined to the Unity module (USA).	3
1.2	Focal length calibration target array for the CSVS on board the US Space Shuttle.	4
1.3	Calibration objects used by various authors.	18
2.1	Overall image formation block diagram.	22
2.2	The Thin lens model.	24
2.3	Lens configuration showing aperture and field stops.	25
2.4	Spherical aberrations in a simple lens.	27
2.5	Off-axis aberrations, coma.	28
2.6	Curvature of the field.	29
2.7	Tangential and sagittal field curvature.	29
2.8	Field curvature and astigmatism in a simple lens.	30
2.9	Axial color in a simple lens (a) and in an achromat lens(b) of identical focal length and speed.	31
2.10	a) Simple lens with minimal principal ray refraction has minimal lateral color. b) Eyepiece with substantial non-symmetrical refraction of the principal ray is afflicted with lateral color.	32
2.11	Idealized CCD and Clock Layout.	34
2.12	CCD spectral responsivity.	37

2.13	Subpixel edge detection based on linear interpolation. (From [43]) . . .	40
3.1	Pinhole camera geometry.	42
3.2	Centroid biases caused by lens distortions.	47
3.3	Centroid location based on two coordinate systems in the camera image plane.	48
4.1	Graphical representation of the calibration object construction tolerances.	54
4.2	Video thresholding for blob detection.	55
4.3	Centroid shift due to bi-linear interpolation.	57
4.4	Centroid shift versus target radius.	58
4.5	Standard centroid measurement error from [31].	59
4.6	Example of camera calibration patterns.	60
4.7	Simulated view from the mirror alignment technique.	63
4.8	Calibration board alignment error (case 1).	63
4.9	Calibration board alignment error (case 2).	64
4.10	Camera calibration rig used for this study.	71
4.11	Calibration board used for this study.	72
5.1	Barycentric local coordinates.	83
5.2	Variation of the gain ratios from the selection of different distorted target elements for the Cohu camera with an 8 mm lens.	85
5.3	Radial distortion effects.	90
5.4	Tangential distortion effects.	91
6.1	Dot locations for residuals calculation.	99
6.2	Camera-space error (CSE).	100
6.3	Example of a volumetric target array.	103
6.4	Volumetric target array configuration.	104
6.5	$RPSE$ arrays definition.	110

A.1	CTVC camera.	133
A.2	ITVC camera.	135
A.3	Cohu camera and Cosmical lens.	137
B.1	Least-Squares minimization flow chart.	142

LIST OF TABLES

2.1	Level of aberrations as a function of aperture and image height. Source [60]	33
4.1	Fundamental camera characteristics.	70
6.1	Ideal calibration dot diameters.	102
6.2	Estimated centroid error in the calibration board.	102
6.3	Ideal dot diameters in the volumetric array.	105
6.4	Estimated centroid errors in the volumetric array.	105
6.5	<i>OSE</i> in the volumetric array at the center of the field of view.	107
6.6	<i>OSE</i> for target at the extremity of the field-of-view.	108
7.1	Calibration performance from the <i>CSE</i> on the calibration board.	112
7.2	Calibration performance from the <i>CSE</i> on the volumetric array.	113
7.3	Calibration performance from the <i>OSE</i> on the volumetric array.	114
7.4	Calibration performance from the <i>RPSE</i> (translation) on the volumetric array.	115
7.5	Calibration performance from the <i>RPSE</i> (rotation) on the volumetric array.	115
7.6	Summary of <i>CSE</i> results on the calibration board.	116
7.7	Summary of <i>CSE</i> results on the volumetric array.	116
7.8	Generic methods gain ratio test results.	117

7.9	Parametric methods gain ratio test results.	118
8.1	Summary of <i>CSE</i> results on the calibration board.	122
8.2	Summary of <i>CSE</i> results on the volumetric array.	122
A.1	CTVC imager properties.	134
A.2	CTVC lens properties.	134
A.3	CTVC camera configuration.	134
A.4	ITVC imager properties.	135
A.5	ITVC lens properties.	136
A.6	ITVC camera configuration.	136
A.7	Cohu camera imager properties.	136
A.8	Cosmicar 6 mm lens properties.	137
A.9	Cosmicar 8 mm lens properties.	137
A.10	Cohu camera configuration.	137
C.1	Calibration performance from the <i>CSE</i> on the calibration board for the CTVC.	144
C.2	Calibration performance from the <i>CSE</i> on the calibration board for the ITVC. (DNC: calibration did not converge)	144
C.3	Calibration performance from the <i>CSE</i> on the calibration board for the Cohu with a 6 mm lens.	145
C.4	Calibration performance from the <i>CSE</i> on the calibration board for the Cohu with an 8 mm lens.	145
C.5	Calibration performance from the <i>CSE</i> on the volumetric array for the CTVC.	146
C.6	Calibration performance from the <i>CSE</i> on the volumetric array for the ITVC. (DNC: calibration did not converge)	146
C.7	Calibration performance from the <i>CSE</i> on the volumetric array for the Cohu with a 6 mm lens.	147

C.8	Calibration performance from the <i>CSE</i> on the volumetric array for the Cohu with an 8 mm lens.	147
C.9	Calibration performance from the <i>OSE</i> on the volumetric array for the CTVC.	148
C.10	Calibration performance from the <i>OSE</i> on the volumetric array for the ITVC. (DNC: calibration did not converge)	148
C.11	Calibration performance from the <i>OSE</i> on the volumetric array for the Cohu with a 6 mm lens.	149
C.12	Calibration performance from the <i>OSE</i> on the volumetric array for the Cohu with an 8 mm lens.	149
C.13	Calibration performance from the <i>RPSE</i> on the volumetric array for the CTVC.	150
C.14	Calibration performance from the <i>RPSE</i> on the volumetric array for the ITVC. (DNC: calibration did not converge)	151
C.15	Calibration performance from the <i>RPSE</i> on the volumetric array for the Cohu with a 6 mm lens.	152
C.16	Calibration performance from the <i>RPSE</i> on the volumetric array for the Cohu with an 8 mm lens.	153

CHAPTER 1

Introduction

The main intent of this thesis is to provide, for practitioners in any scientific field, a reference document containing guidelines for accurately calibrating monocular CCD¹ cameras. Accordingly, the value of this document resides in the fact that it contains practical guidelines for the implementation of the most common camera calibration methods to date, a thorough evaluation of the relative accuracy of those techniques, practical guidelines for the design and fabrication of accurate and dependable camera calibration rigs, a thorough review of the selection criteria for choosing cameras and lenses for vision systems, and finally, a thorough review of external factors to consider in order to generate accurate camera calibrations.

1. Background

Advances in the field of Computer Vision have led to the development of several important technologies primarily in the fields of *quality control* (quality assurance) and *production automation*. Other areas include 2D and 3D measurements, robot vision, identification, printing control and code reading [33]. A great example is a vision system conceived by the National Research Council of Canada (NRCC) in 1971 for highway-barrier studies. The experiment was comprised of a planar array of target dots installed and surveyed in the background of the vehicle-barrier crashing site. The target array was used to geometrically calibrate a high-speed 16 mm camera filming severe automobile crashes. The films were

¹Note that acronyms are defined in the Glossary section.

then used by engineers to assess and improve the energy absorption capacity of different highway-barrier designs.

The capabilities of this system were expanded in 1975 to include the creation of the Space Vision System (SVS). The SVS is a real-time camera-based photogrammetric system which tracks strategically positioned targets on flight payloads, to provide cues for space robotic operations. The system is currently being used by the US Space Shuttle crew members for the assembly and berthing of payloads. The transition from a ground-based application to a space certified system was spread over many years during which the performance, the features and robustness to various lighting conditions were enhanced. The system specializes in “mid-range” photogrammetry where the objects of interest are located in the range of 5 to 25 meters away from the camera.

Since 1989 the technical development of a new version, the Canadian Space Vision System (CSVS), falls under the responsibility of Neptec Design Group Inc.², an engineering firm located in Kanata, Ontario, Canada. The current version of the product is a fully integrated system which processes the NTSC balanced video signal coming from the Space Shuttle Closed-Circuit Television (CCTV) system and presents, for the astronauts operating the robotic arm, an intuitive digital graphics display containing the relative six degree-of-freedom transformation between two payloads. For a comprehensive description of the CSVS, consult [36, 39]. Also note that more pertinent details about the system are given throughout this thesis.

The last mission on which the CSVS played an important role was the STS-88 flight in December 1998. This mission was the first International Space Station (ISS) assembly mission. During this flight, as one of the many activities, the Russian FGB (Zarya) module was installed at one extremity of the US Node-1 (Unity) module. Unity constitutes the core module since most of the principal components of the ISS will attach to it. Zarya is a functional module, equipped with attitude and altitude control thrusters, mainly used before the arrival of the Russian Service Module scheduled for later in the year 2000. The two modules, launched separately, were put together with the Canadian built robotic arm (Canadarm) by astronauts on board the Space shuttle. Figure 1.1 presents a view from one of the Space Shuttle cargo bay cameras taken after the assembly task. During the mission,

²url: <http://www.neptec.com>

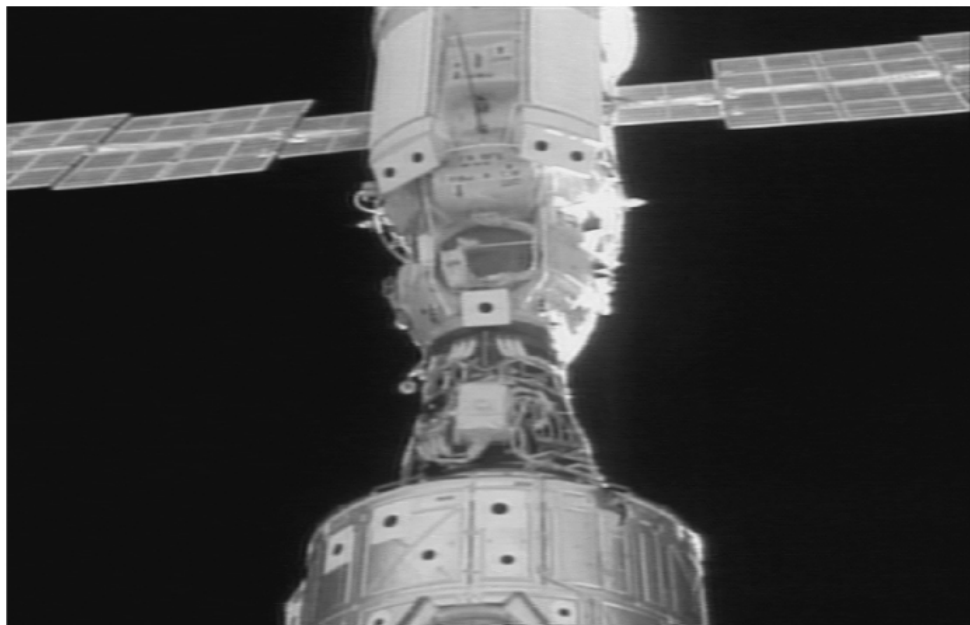


FIGURE 1.1. Scene from the first ISS assembly mission (STS-88). The Zarya module (Russia) had just been joined to the Unity module (USA).

for the last meter of the berthing operation, CSVS was used to track targets (black dots) on the two modules to generate a relative six degree-of-freedom vector between Zarya (top module) and Unity (bottom module), to guide the Canadarm operator.

All cameras used for CSVS operations are calibrated to minimize the non-desirable camera/lens distortions and to specify other critical camera parameters necessary to generate accurate photogrammetric solutions. Prior to a mission, specified flight cameras are characterized in a controlled laboratory environment to measure important parameters that are introduced in the CSVS operation database used during a mission. The choice of cameras, together with their precise adjustments, are determined on a mission-by-mission basis based on the camera locations and on the various target arrays installed on the different modules.

Since most cameras used with CSVS are equipped with automated zoom lenses and are mounted on pan-tilt units (PTU), a special target array has been installed at the far end of the Space Shuttle cargo bay to adjust the camera zoom before every operation. The location of every target of this reference array is precisely known with respect to the center

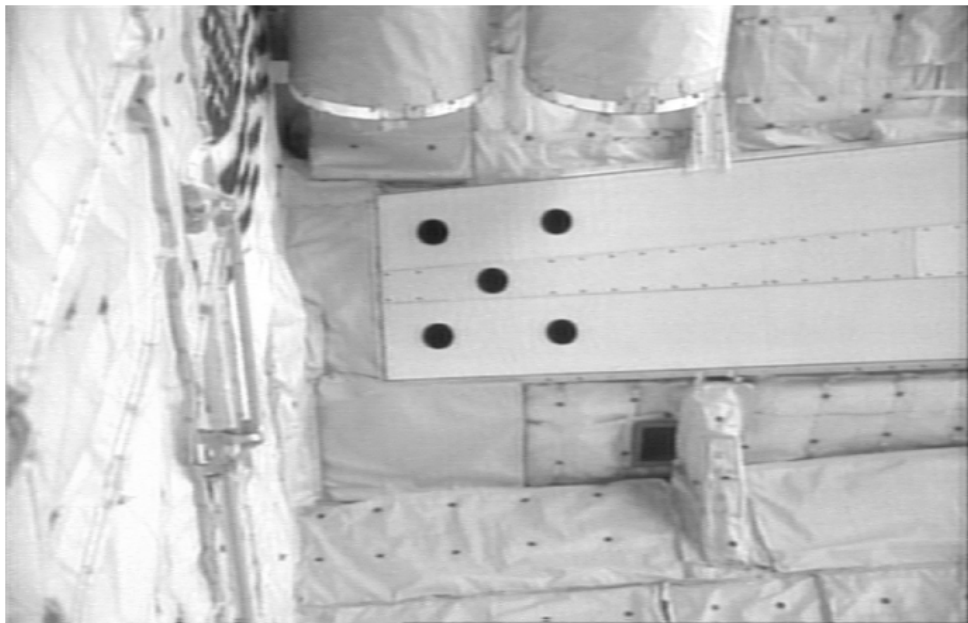


FIGURE 1.2. Focal length calibration target array for the CSVS on board the US Space Shuttle.

of rotation of the PTU. Figure 1.2 was taken during STS-88 while the precise adjustment of the camera zoom was taking place.

A need for the CSVS to have more precise camera characterization data and a better camera calibration apparatus was identified prior to STS-88 during the system certification activities conducted at the David Florida Laboratories, CSA³ in November of 1998. Experimental results obtained at that time led to the conclusion that the CSVS requires very precise calibration information from the flight cameras in order to compute a stable and accurate solution that satisfies tight capture envelopes specified for the STS-88 berthing operation. As a result, Neptec proposed to NASA a thorough investigation to determine how camera calibration could be improved to maximize the CSVS accuracy in its current context of application.

The proposal to NASA was to study new and improved camera calibration techniques, and to establish the requirements for a new camera calibration fixture. A conceptual design to best implement these requirements was also part of the proposal. The results of this investigation are the heart of this thesis.

³Canadian Space Agency

The approach taken for this task was to first assess the current deficiencies in the present calibration fixture and associated calibration method, and secondly to evaluate different camera characterization techniques as published in technical literature. The techniques had to be customizable for the CSVS architecture and had to provide the necessary parameters to generate photogrammetric solutions. This approach was advantageous since a comparison between the various techniques could be obtained using a common vision system.

2. Problem definition

The camera calibration problem revolves around one question: *How do you rationalize and establish a general method to determine the target viewing geometry from the scenes measured by the camera?* Typically, for this kind of problem the solution process goes as follows:

1. Pick a legitimate physical standard (a calibration object).
2. Measure the systems response to the standard (the raw data map).
3. Derive a conversion or correction method to relate the “as measured” values to the “standard” values. This data treatment may or may not have a basis in the physical principles related to the instrument (mathematically manipulated and represented calibration data).
4. Verify that you have generated the proper correction model by applying the correction to the original measurement conditions to check for self-consistency.
5. Validate the correction on an independent standard and confirm that the results are as expected.

These five points define the main research areas related to the problem of camera calibration and also outline the general structure of this thesis.

2.1. The camera calibration challenges. In most computer vision algorithms the formation of images in cameras is described by the *pinhole* model which results from the theories of Projective Geometry [20]. The advantage of the pinhole model is its simplicity and the fact that it can be fully specified by the camera’s *focal length*, *principal point*, and *aspect ratio*. However, this model is only correct if the perspective projection constraint is preserved, i.e., straight lines must be projected as straight lines on the image plane.

Unfortunately, a multitude of sources (lighting conditions, temperature, camera noise, and lens distortions) will disturb this model and therefore appropriate corrections have to be applied to recover the original pinhole model. The distortions of an imaging system are called *systematic imaging errors*. These errors do not include the *stochastic imaging errors* caused by noises during the imaging process. The calibration exercise is to accurately recover the parameters describing the geometry inside a camera despite the error sources, in order to use it for computer vision applications.

Although frequently confounded, the scientific terms *precision* and *accuracy* have two different meanings. Precision describes the statistical variability of the measurements or parameters estimated in the least-squares adjustment, whereas accuracy determines how close the estimated measurements or parameters match the true values. A general characteristic of vision problems is that their solution accuracy and precision can only be obtained at the cost of manipulating accurate input data. In fact, the task of recovering the pinhole model parameters (i.e. the final solution), including distortion correction parameters, is extremely difficult due to the ill-posedness (in the sense proposed by Hadamard [28]) of the problem. A problem becomes ill-posed when there is insufficient information available to obtain a valid solution, i.e., a solution that is physically correct and robust against noise given the choice of metric spaces. For ill-posed problems, the non-uniqueness and numerical instability can have detrimental effects on the final solution if appropriate dispositions are ignored. More specifically, a camera calibration method has to overcome the following problems in order to obtain stable and accurate results:

1. *Existence* - For every sensed element in the visual scene the calibration data has to produce a valid and stable solution that recreates the ideal pinhole model, even though the calibration data was not recorded for every element of the sensor.
2. *Uniqueness* - A central paradigm of projective vision is that given a 2D projection of a 3D object, the inverse problem has a multitude of solutions. *A priori* information, such as the dimensions of a 3D object, has to be used to constrain the solution. Also, the pinhole model is virtual, in the sense that most of the camera parameters are not measurable with a physical instrument, and therefore different combinations of the same (or additional) parameters can produce identical compounded results without

the user realizing it. For example, the focal length and the image scales cannot be uniquely determined at the same time.

3. *Continuity* - The requirement of continuity specifies that when the error on the input data tends to zero, the induced error on the solution tends to zero. Continuity is also related to stability and robustness. Hence, the effect of bounded noise in the data, which is always present in the measurements, must lead to bounded errors in the final solution.

An additional challenge of the camera calibration problem arises from the fact that the measured parameters could change with environmental conditions and lens settings. Thus, in a real life application, a set of calibration parameters will have a small range of action if the calibration conditions are not closely maintained.

To facilitate convergence to a solution, the problem at hand can be conditioned using multiple static cameras/sensors (e.g. stereo vision) [22, 51] or by applying active vision principles [45, 6, 21]. More detailed background on active vision can be found in the general papers by Aloimonos and Bandyopadhyay in [1], and by Bajcsy in [4]. In short, the advantage of active vision is that complicated nonlinear visual processes can sometimes be converted into linear problems, thereby eliminating the multiple solution problem and improving stability.

For this thesis we have chosen to focus our attention on assessing the performance of various static monocular camera calibration schemes. This choice emerges from the fact that the optimization of the fundamental problem could, in turn, support the improvement of more practical and advanced extensions such as dealing with zoom lenses, active vision systems, and the combination of multiple sensors.

3. The state of the art

Photogrammetrists were the first to be concerned about the accurate calibration of cameras. Since the publication of the first edition in 1944, the *Manual of Photogrammetry* [2] has instructed many scientists on numerous subjects related to the recovery of camera parameters, including the correction of undesirable lens distortions. Although the methods were first addressed for photographic film cameras, some of the concepts are still relevant today for performing accurate camera calibrations. Nevertheless, this research area is still

very active today due to the evolution of CCD camera and lens technologies which has triggered the need to adapt and refine old processes. Judging by the number and the diversity of recent articles on camera calibration, no single solution, or group of solutions is being considered “state-of-the-art”. Rather, techniques are chosen, implemented, and perfected based on the user’s needs and budget. For instance, factors like simplicity, calibration rig cost, desired accuracy, speed, and autonomy will dictate the most favorable calibration method for a given application.

The various monocular camera calibration techniques studied in this thesis are grouped into two main categories: *generic methods* and *parametric methods*. A brief review of their principal characteristics is presented in the following subsections. Different approaches have also been proposed to compare different techniques and a brief discussion of those concepts is also included.

3.1. The parametric calibration techniques. A parametric calibration method consists of determining and solving a set of nonlinear equations dependent on the knowledge of the physical interaction between a camera, a lens and features in the 3D world. The main advantage is that there is no approximation involved and the model can be quite elaborate. The problem arises with the complexity of performing the nonlinear optimization and the need for a good initial guess.

Early, scientists like Brown [11] and Faig [19], introduced parametric models of lens distortions for the calibration of film cameras. Their methodology was based on geometrical considerations, to which a mathematical model had been added. The task was to solve the “calibration problem” by relating abstract model parameters to the physical observations. This process was called *bundle adjustment with self-calibration*. Its mathematical formulation is based on the photogrammetric collinearity equations, in some cases augmented with functions to model the geometrical deviations of the physical reality from an ideal imaging system. The collinearity equations are presented later in Chapter 3, however we can say that they express the perspective projection concept of the pinhole camera model.

Later, scientists in the fields of Robotics and Computer Vision, redefined the models and the calibration methodology for vidicon and CCD cameras. A well-known example is Tsai in [55]. In subsequent years, new methodologies for calculating internal camera parameters were proposed by Tsai and others. Such new methods included the precise measurement of

the horizontal scale factor and the optical center [35, 68, 67] to be used with a radial lens distortion model. The main advantage of this work was to limit the number of parameters to be calculated in the global minimization, or to provide a good initial guess. The model was also refined by adding terms for tangential and thin prism distortions [65, 53] present in lenses on some occasions. The thin prism has the particularity of adding radial and tangential distortions, a phenomena which was ignored by the radial parametrization. More recently, researchers have focused on the problem of over-parametrization which can lead to a degradation of the final results.

The following paragraphs present, in chronological order, more details of some important contributions from various authors to improve the accuracy obtained with parametric monocular camera calibration methods.

3.1.1. *Tsai, 1986-87.* As presented earlier, the papers by Tsai [55, 56] introduced a simple, close-form approach for performing parametric CCD camera calibrations. This concept has since been adopted by many scientists in this field. The method is based on recovering all the camera parameters subject to the radial distortion effects through two sequential stages:

- *Stage 1* - involves computing the 3D orientation and the position (distance from the camera to the target excluded) of the camera with respect to the object-world coordinate system.
- *Stage 2* - methods are introduced to calculate the effective focal length, the range distance and the coefficient of a first order radial distortion model.

Stage 1 ignores the lens distortions and treats the coefficient of the perspective transformation matrix as unknowns to be solved by linear optimization methods. The sets of equations are formed given the 3D locations of a series of points in a plane and their corresponding 2D image coordinates. Using the resulting five degree-of-freedom camera location, the rest of the camera parameters are computed in Stage 2.

3.1.2. *Faugeras and Toscani, 1987.* Presented in [22] is a closed-form calibration technique which computes the camera parameters directly from analytical formulas. The method first estimates, from mathematical constraints and camera physical considerations, the linearized global camera parameters. In the second step, the residual lens distortions

are locally corrected throughout the image plane using patches formed with the calibration points. This method takes advantage of the Kalman filtering properties which can provide a measure of uncertainty for the camera parameters given the estimated noise of the input data. This method requires multiple plane calibrations, making it necessary to move the calibration target, or the camera, to different locations using accurate translation stages.

3.1.3. *Puget and Skordas, 1990.* Puget and Skordas in [45] discovered that the method from Faugeras and Toscani returned different homologous intrinsic camera parameters depending on the camera position used during the calibrations. However, the method was very accurate in the computation of the perspective matrix for a given camera location. The approach proposed by Puget and Skordas to solve the parameter uniqueness problem was to repeat Faugeras and Toscani's calibration technique at N different locations and then, using the perspective transformation matrices from each set, apply a minimization criterion leading to the computation of an optimal set of intrinsic parameters. Using the formerly computed intrinsic parameters for each calibration position, the camera rotation and translation were computed by minimizing a mean square error function.

3.1.4. *Lenz and Tsai, 1988.* The paper from Lenz and Tsai [35] makes two contributions. First, it provides a new approach for calibrating the horizontal scale factor of CCD sensors using a one-dimensional Fast-Fourier Transform (FFT). Second, it introduces new methods for measuring the image center, i.e., the intersection of the optical axis with the camera sensor.

The conventional approach for obtaining the scale factor is to refer to the size specified by the CCD manufacturer. However, this usually yields inaccurate results since the timing for the scanning and the digitization is not synchronized. In systems with Phase-Lock Loops (PPL), imprecision due to the computation of pixel size from the sensor element spacing and the sensor clock, as well as inaccuracies in sampling clock frequencies, introduces an apparent change of CCD horizontal size. Tsai and Lenz's concept for accurately measuring the CCD sensor size is founded on the knowledge that the analog output from conventional cameras is formed by superimposing the ideal output value representing the image intensity, some noise, and small signal spikes of the same frequency as the pel shift clock. The spatial frequency of these stripes can be measured and the result corresponds to the frequency difference between the pel clock frequency of the camera and the sampling frequency of the

A/D converter. In most cases, more accurate results were obtained by Lenz and Tsai using this method rather than just using the specifications provided by the manufacturer.

In his earlier papers, Tsai [55, 56] had assumed that the image center was known *a priori* and was considered to be at the center pixel of the CCD array. From simple analysis, it was clear that this selection was not suitable for the radial distortion model since the origin of distortions is not generally co-located with the center pixel of the CCD array. Three groups of methods for measuring a more accurate image center are outlined in [35] and are presented below.

- The *Direct Optical Method* is done independently of all other camera calibration parameters and requires a laser mounted on a four degree-of-freedom adjustment bench. The laser is directly aimed towards the center of the lens making sure all reflections out of the lens coincide with the primary beam, which then indicates that the lens optical axis has been found. This method is similar to auto-collimation techniques, as it uses the reflection of optical surfaces as an alignment cue.
- The *Method of Varying Focal length* is reserved for zoom lenses. When the zoom is varied on a lens, only one point in the image, the center of zoom, will appear to be stationary. The center of zoom is then assumed to be at the intersection of the optical axis with the image plane, which in turn is the image center.
- The *Radial Alignment Method and the Model Fit Technique* uses a physical setup where the camera is pointing towards a photographic glass plate with a coplanar set of calibration points whose coordinates are known. The test pattern is used to fit a simple function based on the Radial Alignment Constraint introduced in [55].

3.1.5. *Beyer, 1992.* In [7], Beyer revisits the fundamental concepts of *bundle adjustment with self-calibration* to include a new coefficient characterizing the *shear* effects that are included when composite video signals and Phase-Lock-Loops (PPL's) are used together. It helps correct for repeatable instabilities in the line-synchronization.

According to Beyer, not all parameters can (or should) be determined during a calibration since over-parametrization can lead to a degradation of the final results. The selection of a particular parameter set, for lens or CCD distortions, depends on the application and the accuracy required. To be completely consistent, Beyer suggests that methods should

be used to evaluate the accuracy and the determinability of additional parameters prior to acquiring the imagery. However, no specific method is offered nor suggested by the author.

3.1.6. *Nomura, Sagara, Naruse and Ide, 1992.* An interesting camera calibration concept was introduced by Nomura *et al.* in [40]. The parametric camera calibration is simplified by decomposing the problem into sequential vertical and horizontal directional calibrations. With this approach, the number of parameters to be calibrated at one time is drastically decreased from 12 to 5 and 4 respectively. The decoupling of vertical and horizontal parameters is possible by using the point symmetry characteristic of image distortions. The 2D model fitting is then decomposed into two 1D fittings using the column and row across an image distortion center. In addition, some parameters are eliminated by the precise knowledge of the location of the calibration chart.

3.1.7. *Tarel and Gagalowicz, 1994.* In [52], Tarel adapts the calibration model introduced by Tsai, to take advantage of the geometric properties obtained by the perspective transformation of ellipse targets on the image plane. The idea combines target centroids and first and second moments of elliptical shapes, which are then fed into the camera parameters estimation algorithm. By using this additional data from the sensed target elements, Tarel reports an increase in the final calibration accuracy.

3.1.8. *Willson, 1994.* In [67] the calibration of the camera parameters, including a first order radial distortion coefficient, is done following a two-step approach similar to Tsai's method. However, the parameters of the first step are used during the second step as initial conditions for an iterative non-linear optimization which refines the complete set of camera parameters. A modified Levenberg-Marquardt algorithm and a finite-difference Jacobian are employed to minimize an error function for all points in the calibration array.

Included in Willson's research is a comprehensive taxonomy that includes fifteen different definitions of the image center and procedures for measuring it. These concepts have also been presented by Willson in [68].

3.1.9. *Tsatsakis, Kayafas, Loumos and Cambourakis, 1995.* In [57], the radial and decentering distortion models are extended to include applications where the lens focus is set to finite distances. These models are based on the theories by Fryer [23] which express the change of distortion coefficients with the defocusing of lenses.

3.1.10. *Shih, Hung and Lin, 1996.* In [48], it was stated that nonlinear methods which estimate all the camera parameters in the same minimization usually produce superior calibration results than Tsai's standard method. However, the calibration of the orientation parameters is more sensitive to noise. For that reason, Shih *et al.* worked on an error sensitivity analysis for the different parameters (intrinsic and extrinsic) involved in the global minimization, to determine which parameters require good initial guesses, and which should be computed separately. The ultimate goal was to establish the best methodology to solve for the unknown camera parameters. They concluded that it was crucial, in terms of robustness, to have a very good *a priori* estimate of the image center before attempting the general non-linear minimization. Otherwise, initial errors in the image center location are easily redistributed among other calibration parameters during the minimization. The covariance and sensitivity equations for the rest of the camera parameters are also provided. These equations are expressed analytically as functions of the effective focal length, CCD sensor area, size of one photo sensor cell, average object distance and the relative object depth.

3.1.11. *Zhou, Tang and Yuang, 1996.* With the same idea of optimizing the methodology to obtain better camera calibrations, Zhou *et al.* [69] proposed a method to help in selecting the distortion models together with their level of complexity. The selection criteria, based on a Student's t-test, is dependent on both the signal-to-noise ratio of the input data and the magnitude of the distortion correction parameters. The authors come to the conclusion that strategic choices have to be made about the parameters to be calibrated in order to maximize the accuracy of the global camera calibration.

3.1.12. *Similarities between the parametric methods.* The parametric calibration methods just described have the common task of uncovering a minimum of eleven coefficients (six extrinsic and five intrinsic) that fully parameterize the pinhole camera model, including the correction of distortions. The difficulty of the task arises when trying to find these parameters through computing the inverse function due to the ill-posed nature of this problem.

The contributions of the authors cited above was to implement methods to improve the stability and the final accuracy when solving the *bundle adjustment with self-calibration* problem. Such solutions include starting with good initial guesses, breaking down the

problem into simpler ones, and better selection of the number of parameters that are solved simultaneously.

3.2. The generic calibration techniques. The generic methods are based on another approach, much less known and less used than the parametric methods, dedicated to solving the camera calibration problem. With the generic calibration methods the explicit expression of the distortions is not assumed. Also, they do not necessarily aim at computing all the extrinsic and intrinsic camera parameters, since once the distortions are handled, linear approaches can perform the task very well. Despite this apparent simplicity relative to the parametric methods, some generic methods claim superior performance. We will now briefly cover, in chronological order, the most notable ones.

3.2.1. *Martins, Birk and Kelley, 1981.* The *two-plane* camera calibration method was first proposed by Martins *et al.* in [38]. The approach is based on measuring the line-of-sight, theoretically for each pixel in the image plane, using two planar calibration arrays. Thus, the calibration process consists of estimating the transformation between image centroids and 3D point coordinates in the two real calibration planes. Hence, for a given pixel element (u, v) , two points $p_1(X_1, Y_1, Z_1)$ and $p_2(X_2, Y_2, Z_2)$ on two different planes are defined such that the functions $X_1 = f_{X_1}(u, v), \dots, Z_2 = f_{Z_2}(u, v)$ produce a straight line-of-sight. The calibration consists of finding the six functions f_{X_1}, \dots, f_{Z_2} for each pixel.

Three solutions have been proposed to approximate these functions: (1) a linear transformation, (2) a quadratic transformation, and (3) a linear spline method where triangular patches are built on each calibration plane with vertices located at the calibration points. From their experiments, Martins *et al.* demonstrated that the two-plane spline method provided the best accuracy. As the line-of-sight can hardly be generated for all pixel elements, the model must employ interpolation between the reference points. Since the interpolation generally becomes less accurate as the distance between the reference points is increased, the accuracy of the method is further improved as the grid density is increased.

The main limitation of this method is that it only solves the back-projection problem, i.e., given a pixel in the image, compute the line-of-sight vector through the pixel (or from the pixel to the world). The direct projection problem would be, given the location of a point in space, predict its location in the image (or project the object point into the image).

3.2.2. *Isaguirre, Pu and Summers, 1985.* The two-plane method was refined in [32] to be used for robot/sensors calibrations. The original method was expanded to include the position and orientation of the camera focal point with respect to the target arrays. An iterative robust approach, based on Kalman filtering, is used to estimate the final parameters, together with their uncertainty.

3.2.3. *Gremban, Thorpe and Kanade, 1988.* In [27], while applying the two-plane method, a new procedure was developed to solve both the projection and the back-projection problems. Returning to the pinhole camera model and using an interpolation scheme from the two-plane method, the necessary parameters for photogrammetric applications were retrieved. The effective focal point present in the pinhole camera model is obtained by finding the point in space that minimizes the distance to all the rays. The rest of the camera parameters are easily obtained from subsequent steps once the distortions are corrected for.

3.2.4. *Champleboux, 1992.* In [15], a distortion correction model called N-Plane-Bicubic-Spline (NPBS) was implemented based on the two-plane method. The original theory was extended to include any number of planes. This new method uses the well known Spline theory to approximate the six functions f_{X_1}, \dots, f_{Z_2} associated with each sampled pixel. The results from this paper indicated that the NPBS method produces more accurate results than other two-plane methods to compensate for lens distortions.

3.2.5. *Butler and Pierson, 1991.* Other generic calibration methods, not based on the two-plane formulation, were also proposed. The generic calibration concept proposed by Butler and Pierson in [12] was to use a mosaic of locally affine functions defined between neighboring points in the image plane to correct for lens distortions. The process creates simplex meshing (3 vertices form a simplex) and calculates the correction necessary to recreate the perfect grid. The method uses a flat calibration board with regularly spaced targets or lines, positioned so that it is parallel to the image plane. A similar concept proposed earlier in [26], used an iterative reconstruction approach based on Bezier patches to best-fit the transformation that rearranges the distorted image points to form the ideal image points.

3.2.6. *Peuchot and Saint-André, 1992.* In [44], Peuchot and Saint-André proposed a straightforward generic calibration method that avoids the need for complex empirical formulas to corroborate experimental data. The argument is that so many different error

sources exist in CCD-produced images that one can adopt a general, non biased approach to remove them. This approach considers the camera as a data recorder which is analyzed by comparing the recorded signal and the reference signal. In this case, the only reference is the image of a grid. Only two hypotheses are necessary, first, the undistorted camera projection system follows the pinhole model, and second, the sensed data is considered reliable in the sense that it is stable and repeatable under similar conditions. A reference grid with 45° oblique lines is used as a reference object for the calibration. The corrected reference grid geometry is reconstructed point-by-point from the distorted sensed data simply following the pinhole projection constraint. New points are corrected using a linear interpolation technique in a relatively dense network of points. Note that this method was also adopted by Brand *et al.* in [10], for their research.

3.2.7. *Qiu and Ma, 1995.* A non-parametric approach for camera calibration called *Constrained Topological Mapping* has been introduced in [46]. The prominent concept is an iterative process used to reduce the Euclidean distance between a topological weighted map of points and the input image points. From N input points, a $(N - 1)$ dimensional self-organizing topological map in N -dimensional input space is formed. At the end of the process, a set of correction factors (δ_u, δ_v) can be generated for any point (u, v) in the image plane. The concept of self-organization was borrowed from Kohonen in [34].

3.2.8. *Similarities between the generic methods.* The most accurate and conceptually simple method for camera calibrations would be to measure the calibration parameters at each pixel in the image, i.e., at each pixel measure the line-of-sight vector. This would produce a gigantic lookup table. Then, simple indexing would allow one to search and find the line-of-sight vector that passes closest to the point in question. The obvious compromise is to sample the image and interpolate between a smaller but sufficient number of data points. If the error in interpolation is less than the measurement error, no accuracy is lost.

The main concept that is shared by the generic camera calibration methods is to put aside the *bundle adjustment with self-calibration* approach during the calibration since the parametric lens distortion models are not always verified in practice. This can be done since no *a priori* knowledge about the distortion type is necessary to comply with the perspective constraints. The idea behind most generic techniques is, once the various distortions are corrected, replacing the data in the pinhole model is an unproblematic task.

3.3. Different calibration methods, same goal. The parametric and the generic calibration methods ultimately have the same goal, that is to accurately retrieve the fundamental, undistorted parameters that describe the internal and external geometry of a CCD camera. Accurate information is crucial for applications in Computer Vision and Robotics to solve problems such as *what* a machine may see and *where* an object is from the image taken by the CCD camera and ultimately from the hand of the robot. Only when the camera is properly calibrated can the 2D image information be transformed accurately back into real-world locations.

3.4. Calibration accuracy evaluation techniques. One of the first questions asked about any camera calibration is: *how accurately does it capture the image behavior?* This information is necessary for estimating the performance or accuracy of the camera in the context of its use. From the different calibration methods reviewed in §3.1 and §3.2, two categories were retained for monocular camera calibrations: the *image-space* and the *object-space* evaluation techniques.

3.4.1. *Image-space evaluation techniques.* One of the easiest methods to implement for verifying the accuracy of the camera calibration is to measure the *radius of ambiguity zone*. Given the location of a point in space, the image coordinate obtained from its direct projection onto the image plane is compared to the new corrected image point obtained after applying the camera calibration method. The magnitude of this correction residual is obtained by calculating the root-mean-square from both image axes. Such a method is frequently used in the literature [55, 27, 12, 65, 67, 46].

3.4.2. *Object-space evaluation techniques.* A second category of error metrics for monocular systems is the direct 3D measurement accuracy of a calibrated vision system used in a known 3D environment. The first approach in this category uses a photogrammetric solution to derive the 3D location of known reference points and then compares those results with the 3D survey data. The second approach does not require solving for unknown parameters. For that technique, the centroid of an object feature is projected back through the camera model and then one has to calculate the closest distance of approach between the image point's line-of-sight and the point in 3D space that was supposed to have cast the image.

Implementation examples of such methods can be found in [38, 27, 15, 67].

3.5. The most common calibration objects. The accuracy of the calibration object is crucial for camera calibrations. It serves as “ground truth” for establishing the projective scaling obtained with a given camera. Also, the optical nature of the landmarks on the object must be selected to ensure high signal-to-noise ratio for optimal feature detection. Figure 1.3 presents some common calibration objects encountered in the literature and summarized in [14].

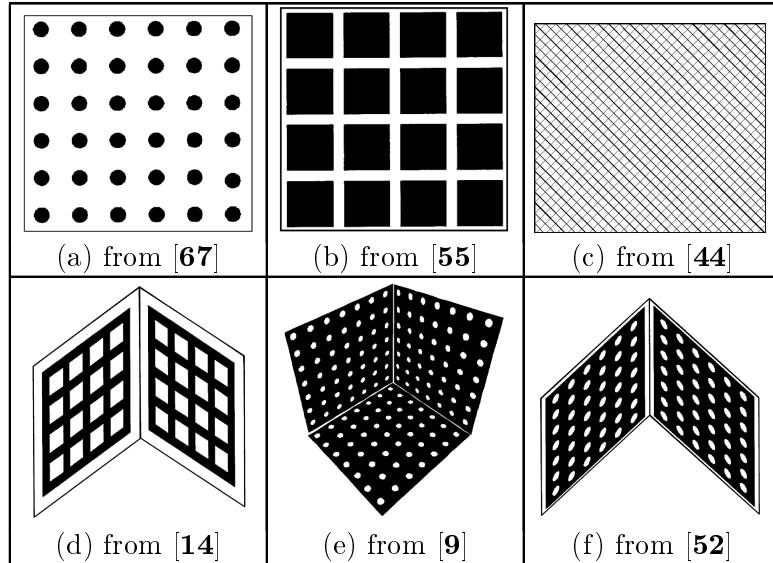


FIGURE 1.3. Calibration objects used by various authors.

In Figure 1.3, calibration objects (a) through (c) are made of planar surfaces with patterns painted, printed or glued on. Sometimes, highly accurate translation and/or rotation stages are used to move the calibration array to different locations. Target objects (d) through (f) are made to instantaneously create an accurate 3D environment. With such objects, the camera is positioned in front of the target and only a single static image sequence is necessary to conduct the calibration.

4. Main contributions

This research makes contributions to computer vision in the selection, refinement, evaluation, and implementation of accurate camera calibration methods, specifically:

- *Practical guidelines for the implementation of the most common camera calibration methods to date.* We have implemented fourteen calibration methods which were

inspired by advanced techniques in the field of geometric camera calibration. In this thesis, we present the various steps and strategies that lead us to the successful implementation of these different methods.

- *A thorough evaluation of the relative accuracy of the most common camera calibration methods to date.* The number of tests that can be used to evaluate the accuracy of the monocular camera calibration is fairly limited. However, we expand this list with a new methodology that looks at the consistency of the calibration parameters by comparing the data obtained from various regions on the image plane. Using these tests, a statistically sound approach is then used to compare the relative performance of the calibration methods tested.
- *Practical guidelines for the design and fabrication of accurate and dependable camera calibration rigs.* While preparing this thesis we gained a significant amount of knowledge about the design, construction and operation of accurate calibration rigs. This experience was first acquired by using a calibration rig developed by the NRCC for calibrating flight cameras for NASA. Second, we had to entirely design and fabricate a new calibration rig for this study. Lastly, our involvement at Neptec Design Group Inc. with the construction of a state-of-the-art camera calibration rig gave us insights to all the aspects necessary to obtain a high quality and functional rig. From that experience, we present a list of factors that need to be considered when building such high-accuracy calibration equipment.
- *A thorough review of the selection criteria for choosing cameras and lenses for vision systems.* Various factors influence, with more or less significance, the precision and accuracy of CCD cameras. For that reason, it is our opinion that cameras used for computer vision applications should not be exclusively considered as “black boxes” employed to capture images. Rather, a minimum of information about the characteristics of the cameras is necessary in order to derive the expected performance of a vision system.
- *A thorough review of external factors to consider in order to generate accurate camera calibrations.* A list of the most important external agents contributing to successful camera calibration is presented. The list includes: lighting control, system hardware/software configuration, calibration laboratory, etc.

5. Thesis outline

The remainder of the thesis is structured as follows:

Chapter 2 introduces the general image formation concept adopted for this thesis. The concept incorporates, in a simplified fashion, the transformation of the input signal (incoming light), from its photon stage when it enters the camera lens to the final digital image format. Understanding the various steps and error sources allows a better understanding of the camera calibration in terms of separating the phenomenon that are *systematic imaging errors* that can and should be corrected and the *stochastic imaging errors* that cannot be corrected, but should be controlled.

In Chapter 3 we review the mathematical formulas that are employed to completely define the pinhole camera model. Those formulas set the basic theory for all camera calibration methods implemented for this thesis.

Chapter 4 discusses the importance of selecting and setting up a proper calibration rig. The equipment must be suitable for all calibration methods and must provide optimal test conditions to justly segregate the different calibration methods based on their true performance. This chapter also presents some common verification tests for CCD cameras, and describes the software/hardware used to perform the data collection.

Chapter 5 contains a description and a methodology for implementing and performing the various calibration methods tested as part of this thesis. The *generic* calibration techniques chosen are presented first, followed by the *parametric* methods.

Chapter 6 describes the common means of measuring the performance of the different calibration methods using the calibration rig itself and a new target array, called the *volumetric array*. The volumetric array is the primary camera calibration evaluation setup used for this study. It is effectively an independent test bench recreating the size and distance from the camera of a typical flight payload. The main advantage of such a target array is the introduction of a new performance measurement metric that verifies the consistency of the camera calibration using different areas of the image plane.

Chapter 7 summarizes the large amount of data gathered during the preparation of this thesis and unveils the methods, or categories of methods, that produced the most accurate results.

Chapter 8 concludes this thesis by presenting a review of some important topics and by summarizing the test results obtained. The very last section summarizes new research areas that we would like to pursue.

In Appendix A we detail the characteristics of the cameras and lenses used for this thesis.

Appendix B contains the details of the Least-Squares minimization techniques used to solve the non-linear parametric calibration equations.

In Appendix C we present numerous tables which contain the compiled test results from the performance analyses of all the cameras and calibration methods.

CHAPTER 2

General image formation concept

It is important to understand the nature and extent of a problem before trying to solve it. Therefore, it is essential, before considering the camera calibration problem, to review the concepts behind standard CCD image formation techniques to identify the physical aspects that are subject to adjustment. The following principles will also help the reader understand the notions discussed later, about controlling the camera calibration environment.

A perfect sensor/lens system is designed to produce a “perfect” image, i.e., light from all points on an object are correctly focused on a corresponding conjugate point on an image. Under most conditions the sensor/lens system fails at this task mainly because of some errors in the precision of the Charged Coupled Device (CCD) camera to convert light into an electrical signal and from various aberrations in the optical system. As presented in Figure 2.1, the acquisition system can be modeled by two group effects which are related to the lens system or the CCD camera and which are successively applied to the original unaffected light source.

For the lens system, the aberrations are divided into two effects: the non-linear distortions and the low-pass filtering. As for the CCD sensor aberrations, they can be split into

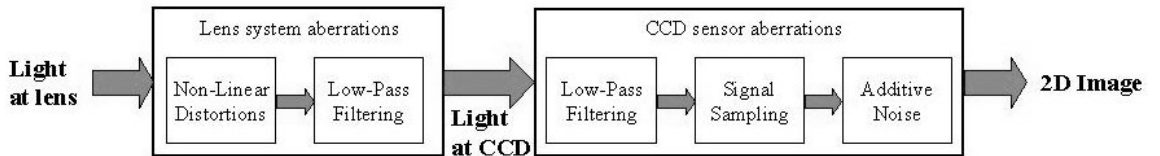


FIGURE 2.1. Overall image formation block diagram.

three effects: the low-pass filtering, the signal sampling, and finally some noise added to the resulting data.

The various aberrations from Figure 2.1 have the effect of moving the mean position of the image features, and blurring the image. Distortion is a unique aberration in that it does not affect the quality of the image in terms of sharpness or focus. Rather, distortion affects the shape of the image, causing it to depart from a true-scale duplicate of the object. Thus, it is possible to compensate for the distortion after the target image is localized. The term *filter* (or *low-pass filter*) is an aberration commonly used to describe a device that discriminates, according to some attribute of the objects applied at its input, what passes through it. The natural effect of filtering in sensing devices is the loss of high-frequency information due to intrinsic system limitations. For example, a high frequency step-like edge in a scene will appear like a gradual, blurred linear region on a photographic film or video signal. Signal sampling is the conversion of a continuous-time signal (analog video for this matter) into a discrete-time signal obtained by taking “samples” of the continuous-time signal at discrete-time instants. Signal sampling is not an aberration by itself but can cause aliasing (loss of high frequency information) and quantization (staircase approximation of the original signal), which are forms of aberrations. By definition, noise is a corrupting signal added onto the original, pure signal. The noise power is assumed to be much smaller than the original signal power. Furthermore, the original signal is assumed to have low frequency components, whereas the noise source is white (zero mean, Gaussian distributed).

The main sources of aberrations from the lens and from the CCD sensor are further described in the following Sections 1 and 2.

1. Lens system aberrations

An ideal optical system is one where every point P in the 3D object-space is imaged as a point p in the image plane, and the image scale is constant. A real pinhole camera exhibits this property very well. However, the level of scene illumination must be relatively high for the sensing medium to receive enough energy. Since they make up for this limitation, camera lenses are most frequently used.

The Thin lens model is commonly use to reduce, by approximations, any complex lens system to a simple pinhole model representation, as shown in Figure 2.2.

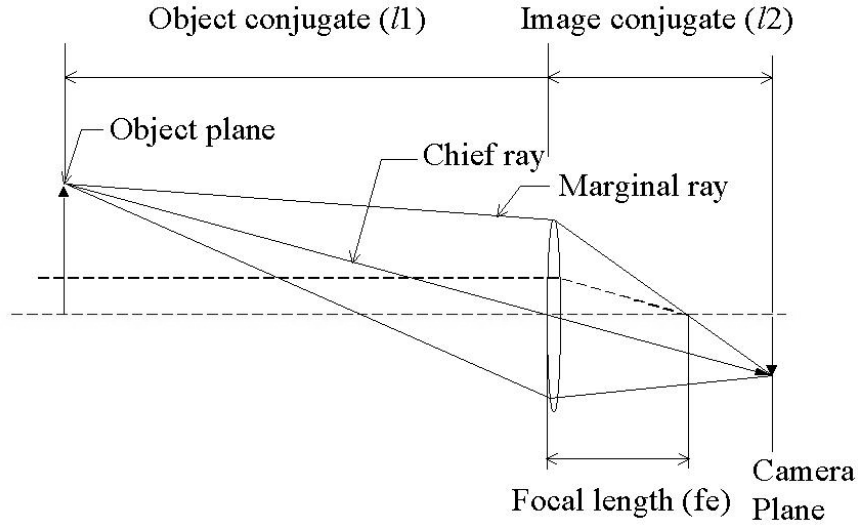


FIGURE 2.2. The Thin lens model.

The distance from the object plane to the lens is called the object conjugate. Likewise, the distance from the lens to the camera plane is called the image conjugate. These conjugates are related by the lens maker's formula given by:

$$\frac{1}{f_e} = \frac{1}{l_1} + \frac{1}{l_2}$$

The focal length is by definition, the distance between the lens and the camera plane when the object is at infinity. Note that if l_1 is equal to infinity then l_2 is equal to f_e . The situation where l_1 is equal to infinity occurs when the lens focus is set to infinity, in which case the lens focal length is correct. All other cases imply that a magnification factor is involved and therefore l_2 varies a few millimeters from the focal length.

As we have said, lenses have their own limitations. Aside from the distortion effect, real lenses always produce a low-pass filter effect on the original image, even when the subject is perfectly in focus. In fact, there is always a bandwidth limitation on the optical lens due to its finite aperture, and therefore it transmits a filtered version of the ideal image on the CCD. Studies show (see [25]) that the low-pass effect is intimately related to the shape of the lens aperture. The mathematical representation of this phenomenon can be obtained through the lens impulse intensity response, known as modulation transfer function (MTF).

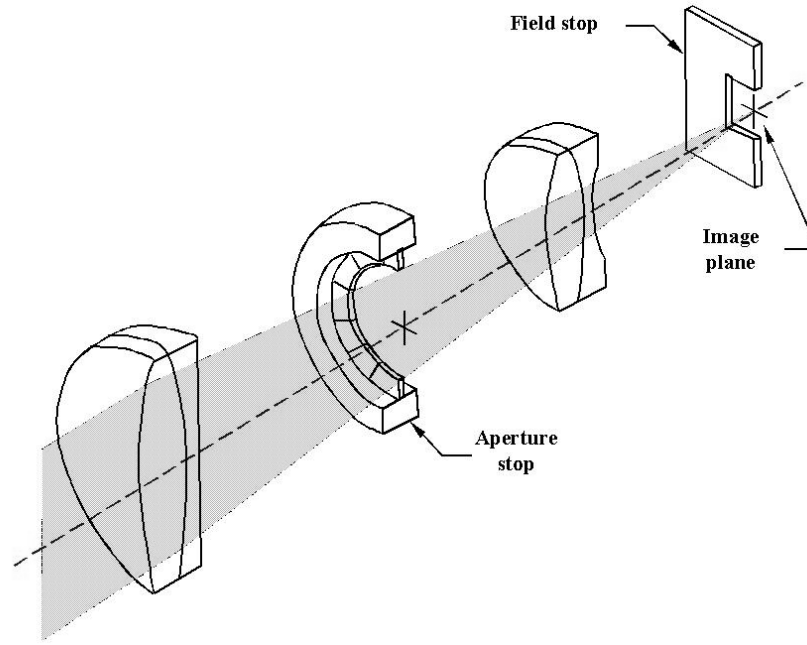


FIGURE 2.3. Lens configuration showing aperture and field stops.

Aperture stop limits the size, and thus the total power, of an incident beam on a collecting lens. The total flux can be varied using an aperture or an iris diaphragm. A field stop controls the size and shape of the image. Figure 2.3 shows a lens system with an aperture stop and a field stop.

There is no interdependency between the aperture stop and the field stop. As the aperture stop is enlarged or reduced, the image size remains constant. If the field stop is enlarged or reduced the image size changes, but the power density in the image remains the same.

The following subsections present the seven most important aberration effects caused by lenses. The effects presented are either *chromatic* or *monochromatic* which refers to the spectral content of the light source. Chromatic light sources include a wide range (bandwidth) of wavelengths and monochromatic light sources are limited to a single wavelength. The mathematical representations of the most significant effects for camera calibrations are presented in Chapter 5, although a more complete list can be derived from [63].

1.1. Geometrical distortions (monochromatic effect). Geometrical distortions from lenses are, by definition, the image point shift from the position predicted through

paraxial approximation (i.e. simple perspective projection onto the image plane modeled by the pinhole camera model). Geometrical distortions will generally preserve the original light content but will shift rays in proportion to the distance from the optical axis and independently of the lens aperture.

1.2. Spherical aberrations (monochromatic effect). Spherical aberration is the imaging error found when a lens is focusing an axial bundle of monochromatic light. Each zone (annulus) of the lens aperture has a slightly different focal length. The further away from the optical axis the ray enters the lens, the nearer to the lens it focuses (crosses the optical axis). This lack of a common focal distance creates blurred circles centered on the optical axis.

The Longitudinal Spherical Aberration (LSA) is measured by the distance on the optical axis between the crossing of the rays which are closer to the optical axis (paraxial rays) and the rays which emerge from the outer section of the lens (marginal rays). The height at which the marginal rays intercept the paraxial image plane is called Traverse Spherical Aberration (TSA). It is possible to analyze the intensity spread function of the light on the camera array and improve the image quality. In almost all applications the quality is optimal when the lens is focused close to the point of minimum spot size; i.e. when the image plane is located somewhere in the middle of the LSA vector (see Figure 2.4).

Spherical aberration is constant over the field of the system and is directly proportional to the lens aperture.

1.3. Coma (monochromatic effect). Coma is an aberration that affects off-axis light bundles in a way similar to which spherical aberration blemishes axial bundles. When an off-axis circular bundle of light is incident on a lens afflicted with coma, each annulus focuses onto the image plane at a slightly different height and with a different spot size. The result is an overall spot that appears as a characteristic comet-like flare, having a bright central core with a triangular shaped flare extending toward the optical axis of the lens. Figure 2.5 illustrates the concept.

Consequently, coma is absent on the axis and it increases linearly with the field angle or aperture.

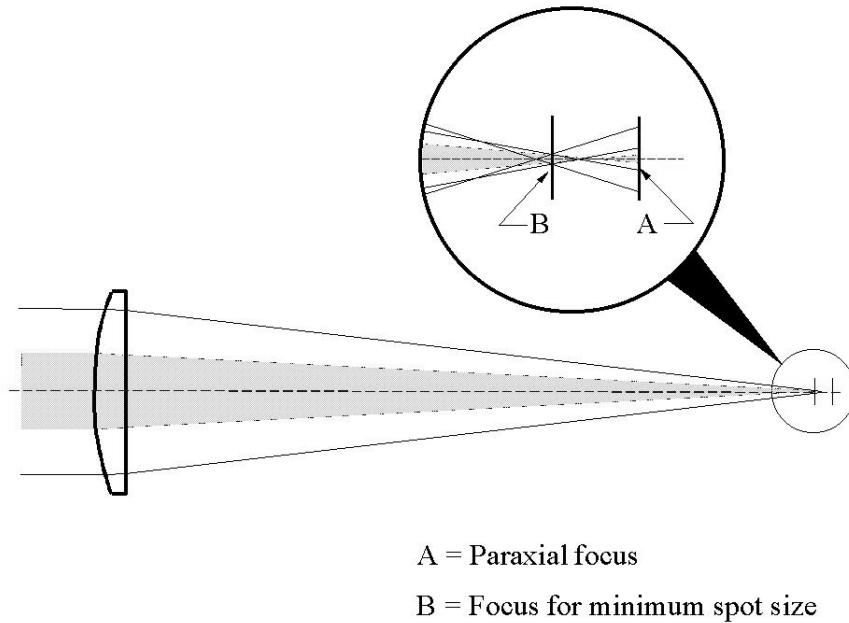


FIGURE 2.4. Spherical aberrations in a simple lens.

1.4. Field curvature (monochromatic effect). Most machine vision systems assume that the image is formed on a planar surface. However, in most optical systems there is a tendency to image better on a curved surface. The nominal curvature plane ($1/\text{radius}$) is referred to as the Petzval or field curvature of a lens (see Figure 2.6). For simple lenses this curvature is approximately equal to the lens power. When the lens is free of other off-axis aberrations, the image is formed on the Petzval surface. The field curvature effect brings the image out of focus the further away from the optical axis the light is projected and the wider the aperture is.

1.5. Astigmatism (monochromatic effect). When astigmatism is present in a lens system, fans of rays of differing orientations at the lens aperture tend to focus on two different curved surfaces.

The curves of Figure 2.7 represent a cross section of the top part of the image surface (paraxial focal plane) from the optical axis out to the edge of the field-of-view. The plane containing both the optical axis and the object point is called the tangential fan. Light rays traveling inside the tangential plane are called tangential rays. The plane that is perpendicular to the tangential fan is called the sagittal fan or radial plane. The principal

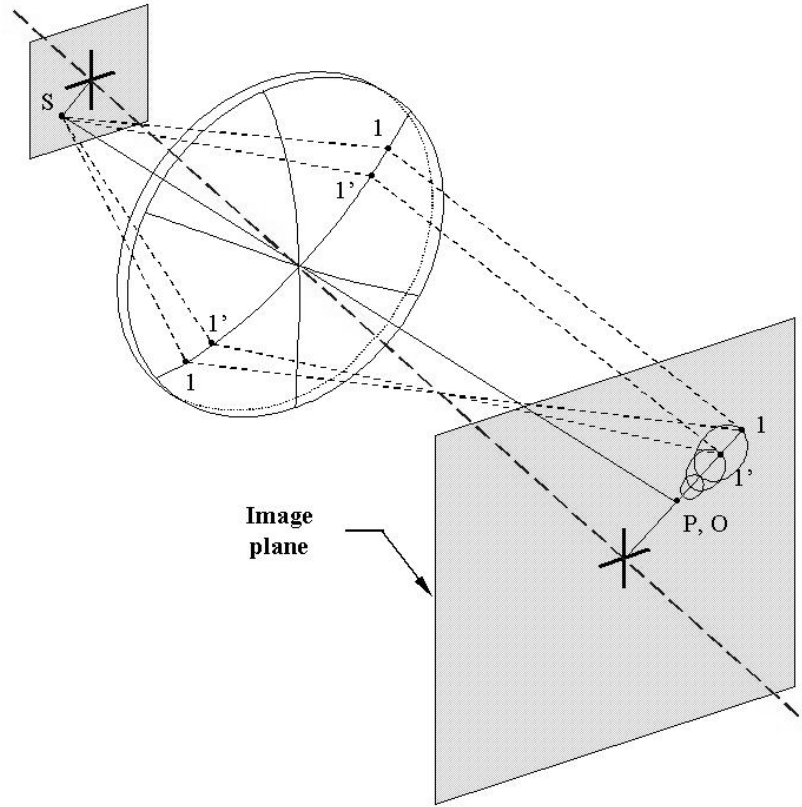


FIGURE 2.5. Off-axis aberrations, coma.

ray goes from the object point through the center of the lens aperture. As illustrated, the tangential rays from the object focus closer to the lens than do the rays in the sagittal fan, which in turn are focused in from the image plane. Clearly, when the image is analyzed at the tangential focus location, a line in the sagittal direction is observed. Conversely, a line in the tangential direction is observed when the image is formed at the sagittal focus location. In between these conjugates, the image is either an elliptical or circular blur. Astigmatism is, by definition, the difference between the tangential and sagittal field curves. If the tangential and sagittal surfaces are coincident, then the lens is free of astigmatism and the image is formed on the Petzval surface.

When astigmatism is present, the tangential field departure from the Petzval surface is three times the departure from the sagittal field, as illustrated in Figure 2.8.

Astigmatism aberrations produce effects similar to the field curvature, and are therefore a function of the field angle and the lens aperture.

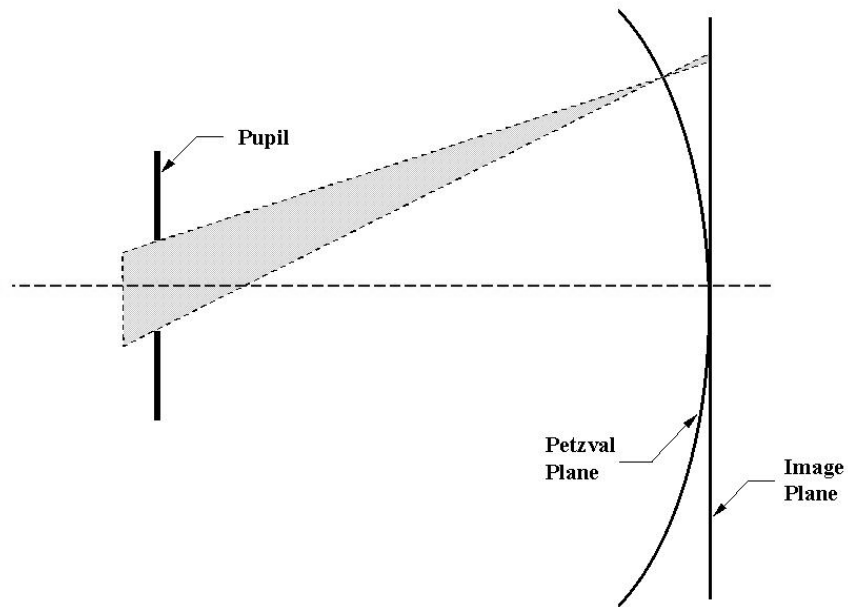


FIGURE 2.6. Curvature of the field.

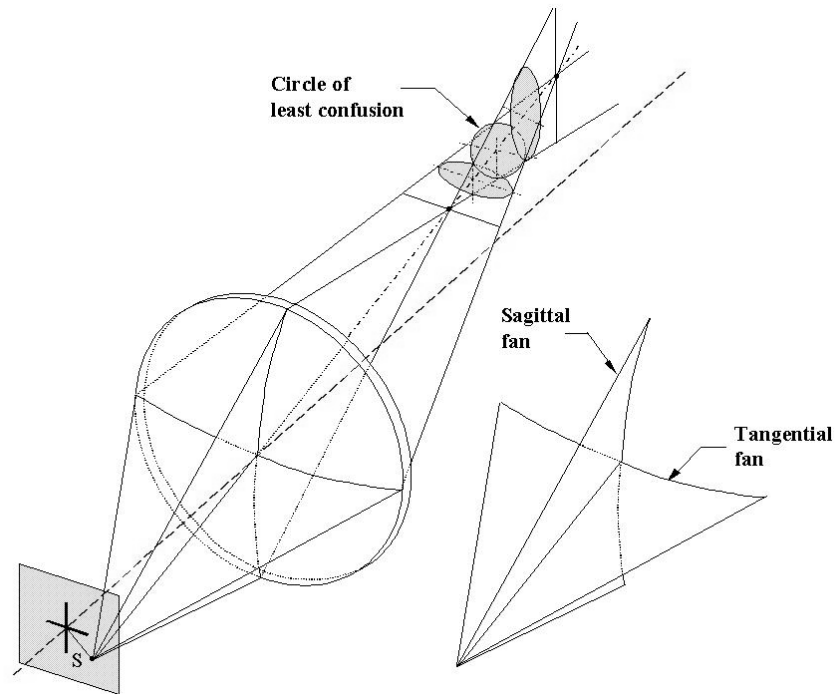


FIGURE 2.7. Tangential and sagittal field curvature.

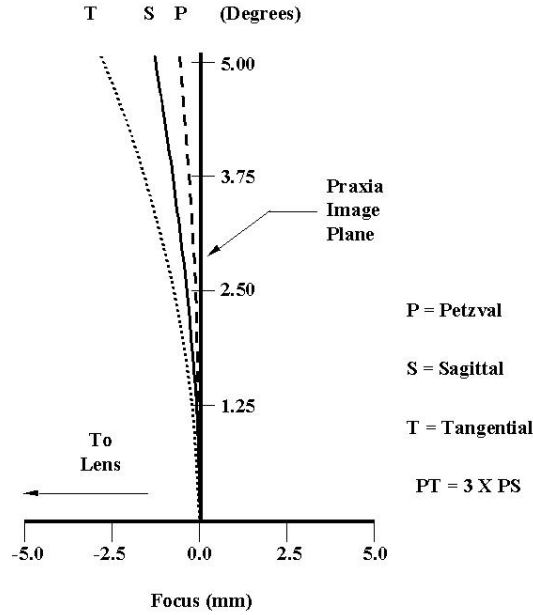


FIGURE 2.8. Field curvature and astigmatism in a simple lens.

1.6. Longitudinal/Axial Color (chromatic effect). For all types of glass, the index of refraction varies as a function of wavelength with a greater index for shorter wavelengths (blue). Also, the rate at which the index changes is greatest at the shorter wavelengths. In a simple lens, this causes each wavelength to focus at a different point along the optical axis. This chromatic spreading of light is known as dispersion and is directly influenced by the lens aperture. Figure 2.9a illustrates a simple lens focusing a bundle of white light covering the spectral band from 450 to 650 nanometers. If the focus is set for the middle of the band, as shown, the blur circle consists of a green central core with a halo of purple (red and blue) surrounding it. Correction for chromatic dispersion can be accomplished by converting the simple lens into an achromatic doublet as shown in Figure 2.9b. The two glass types selected correct the primary axial color by bringing the two extreme wavelengths to a common focus.

1.7. Lateral Color (chromatic effect). For off-axis bundles, the corresponding central ray is called the chief ray, or principal ray. The height of the chief ray at the image plane defines the image size. If lateral color exists in the lens system, the chief ray is dispersed, causing the different wavelengths to be imaged at different heights on the image plane. Because of the change in the refraction index with the wavelength, blue is

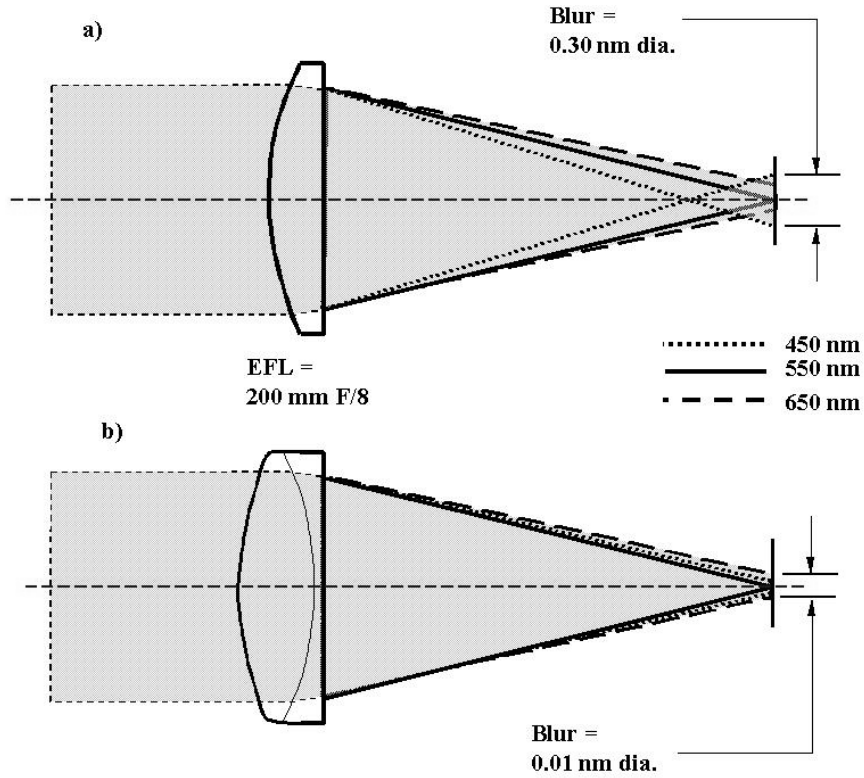


FIGURE 2.9. Axial color in a simple lens (a) and in an achromat lens (b) of identical focal length and speed.

refracted more strongly than red, causing the rays to focus at different heights. The result is chromatic, radial blur for off-axis image points. The lateral color effect is dependent on the lens field-of-view. Figure 2.10 shows the effect of lateral color for a simple lens and for a more complex lens where substantial non symmetrical refraction of the chief ray is present, creating the aberration.

1.8. General lens aberration contributions. Table 2.1 presents the relative sensitivity to lens aperture (Y) and image height (H) for the aberrations previously mentioned. The sensitivity increases proportionally with the order of Y and H .

As described in [49] and [18], many recent technological findings are contributing to the design of low-aberration lenses by including more elements with particular optical properties and shapes. Hence, three methods are particularly proficient at reducing lens aberrations:

- **Aspherical lenses** : aspheric surfaces located strategically in the lens system can significantly reduce spherical and coma aberrations.

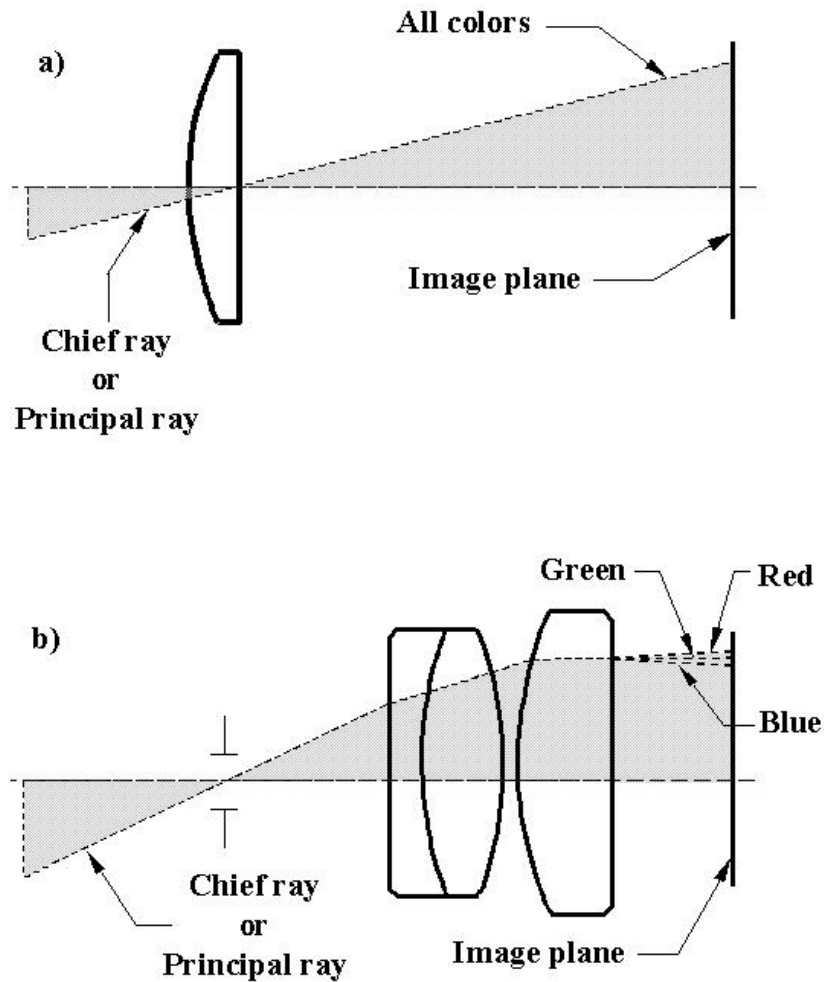


FIGURE 2.10. a) Simple lens with minimal principal ray refraction has minimal lateral color. b) Eyepiece with substantial non-symmetrical refraction of the principal ray is afflicted with lateral color.

- **Symmetrical elements** : a system that is symmetrical about the point where the chief ray crosses the optical axis (aperture stop) has little or no lateral color (or coma) due to the tendency of the aberration to cancel itself as the chief ray traverses the symmetrical halves of the system.
- **Multiple element lenses** : complex lens configurations can minimize the effects of various aberrations such as lateral color (Section 1.7), astigmatism (Section 1.5) or field curvature (Section 1.4). Anti-reflective coatings make this solution possible by

Aberrations	Aperture Y	Image Height H
Spherical	Y^3	—
Coma	Y^2	H
Petzval	Y	H^2
Astigmatism	Y	H^2
Distortion	—	H^2
Axial color	Y	—
Lateral color	—	H

TABLE 2.1. Level of aberrations as a function of aperture and image height.
Source [60]

minimizing internal light dispersion and therefore practically allowing an unlimited number of elements in the optical system.

Therefore in theory, when choosing high quality lenses for computer vision applications, the main systematic aberration worth correcting for is the geometrical distortions.

2. CCD sensor aberrations

The basic operation of the CCD sensor is to convert light into electrons. The light coming from the lens focuses on the image plane (i.e. the CCD sensor surface) and eventually forms an encoded analog signal, which will be sampled over time and space. In practice, a CCD is subject to a variety of physical and performance limitations due to the generation, amplification and formatting of the visual information. Of particular importance to automated imaging systems are those limitations related to the fidelity or quality of the electronic image.

A basic introduction to the Frame Transfer (FT) CCD technology is presented in the following subsections together with the main CCD aberrations. As presented in Figure 2.1, these aberrations can be classified as low-pass filtering, signal sampling and/or additive noise, depending on their nature and CCD operation conditions.

2.1. CCD operation. Following is a somewhat idealized description of the operation of a frame transfer CCD. The basic functional blocks of a typical frame transfer CCD and clock circuitry are shown in Figure 2.11.

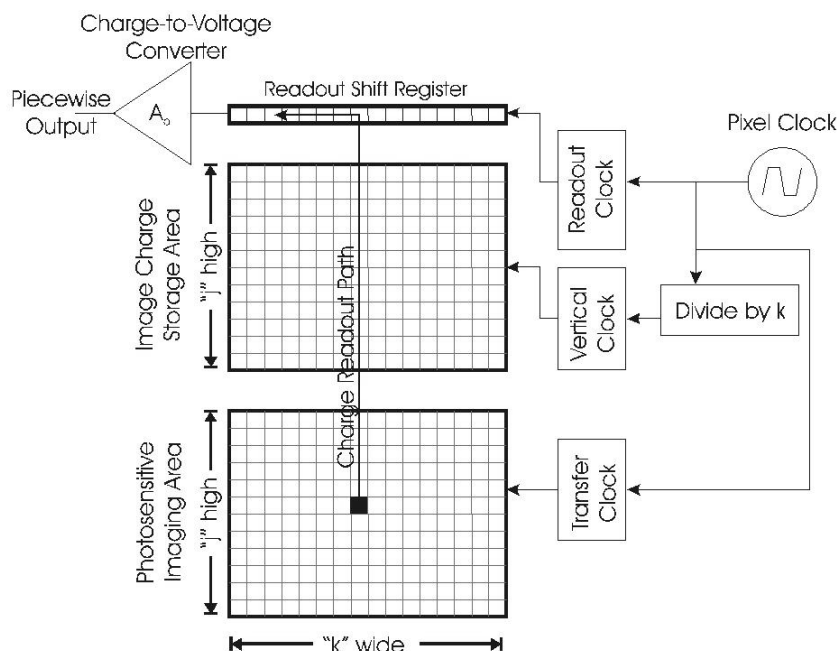


FIGURE 2.11. Idealized CCD and Clock Layout.

The active imaging area of a CCD is usually a rectangular array of photosensitive elements (or pixels). Each of the elements can temporarily store any charge that is liberated by photons interacting within the pixel boundaries. If an image is projected onto the array, the light distribution within the scene produces a corresponding charge distribution within the pixels. By measuring the charge stored in each pixel, an electronic analog version of the image may be generated.

In Frame Transfer CCD's, the accumulated charge in each active pixel is transferred to an intermediate storage cell. Charge transfer is accomplished by shifting the accumulated charge bundles along the corresponding CCD columns in a "bucket brigade" fashion. Frame transfer is done relatively quickly, to purge the last image and allow the photosensitive area to begin acquiring a new image. A multi-output clock generator circuit controls the timing and coordination of this transfer process.

While the next image is being acquired, the intermediate storage cells are read out. This readout progresses akin to a raster scan. Each storage area row is sequentially transferred to the parallel input of a shift register. Between row transfers, the shift register is cycled to

present each pixel's charge to a "charge-to-voltage" conversion stage. This process continues until all of the stored charge has been read.

2.2. Video frame grabbing. For computer vision systems that rely on digital image processing techniques, an additional video digitization step is introduced. The digitization process, during which the analog video signal is converted to discrete intensity measurements, introduces the following effects:

First, the digitizer must define the number of sampled points along each line, and hence the resolution of the image. Because of the variation in clocks between cameras and digitizers, it is common to find distortions of a few percent in pixel shape resulting from the conversion.

Second, the digitizer must control the precision of each intensity measurement. The analog voltage signal is usually digitized with an 8-bit A/D (analog-to-digital) converter; a chip using successive approximation techniques to rapidly sample and measure the voltage in less than 100 nanoseconds, and producing a number value from 0 to 255 that represents the pixel brightness. The quality of most cameras and other associated electronics rarely produces voltages that are free enough from electronic noise to justify more than 8-bit digitization capability. For example, if the number of quantization levels is 256, in a worst case scenario of bilinear interpolation and quantization errors of the same sign, the edge transition is vertically shifted by $\Delta/2$, Δ being the quantization step. Such a vertical transition has to be compared with a total edge transition of approximately 60Δ to 80Δ .

2.3. Spatial resolution. One fundamental performance characteristic of an imager is the effective spatial resolution limit of the CCD. The image projected onto the CCD is effectively sampled at the spacing of adjacent pixels. If the original image contains significant details at frequencies higher than this, aliasing can occur. Normally, the optical projection system is provided with (or may naturally have) a low-pass spatial filter to minimize this problem. The overall effect is that the resulting system does not reproduce high frequency details. There must be at least two samples within the transition time of an edge to avoid aliasing.

Video engineers express resolution in terms of modulation transfer function (MTF). This measure is similar to what is used to assess camera lens resolving power.

2.4. Dynamic resolution. Dynamic resolution is the resolving power of a stationary sensor to fine moving details. A sensor with poor dynamic resolution will show lag in the image. Lag is the inability of the sensor to respond to sudden changes in light intensity. The more lag, the worse the dynamic resolution is, and the larger the limit on object motion is.

2.5. Sensitivity. Sensitivity or responsivity, is a term that describes the efficiency of the CCD to convert light into an electrical signal. The spectral response and the dynamic range are the two components used to describe the sensitivity of a sensor.

The physical process responsible for charge production within a pixel produces a *spectral response* truncated at both short and long wavelengths. At very long wavelengths, the incident photons lack sufficient energy to liberate charge. As the wavelength decreases (the energy increases) and the probability that a free charge will be generated increases. Quantum efficiencies, or the ratio of incident photons that generate charges, typically reach 30% to 70% in the visible range for CCD cameras. As the photon energy continues to increase, the low transmittance of most optical materials at short wavelengths sharply limits the number of photons that actually reach the active region of the CCD. A typical CCD spectral response is shown in Figure 2.12. This characteristic is plotted both in terms of responsivity and quantum efficiency. (Note: the responsivity is defined in terms of the current output produced by a certain optical input power confined to a narrow wavelength band.)

Dynamic range is the measure of usable signal output for a range of lighting conditions, from low light levels to high light levels before saturation. Blooming is a problem that occurs when illumination is so intense that it causes saturation of certain pixels which start to spill charges to adjacent pixels. The result on the video image is bright areas with blurred definition. Smearing is also a phenomenon that happens under extreme lighting conditions. For a frame transfer CCD, the leakage occurs during the frame transfer period. At low-light levels the signal-to-noise ratio becomes very low and makes the video signal unusable. Some of the most frequent CCD noises are described in the following subsections.

2.6. Photon noise. The emission of photons from any source is a random process. The number of photoelectrons collected in a potential well in one integration period is

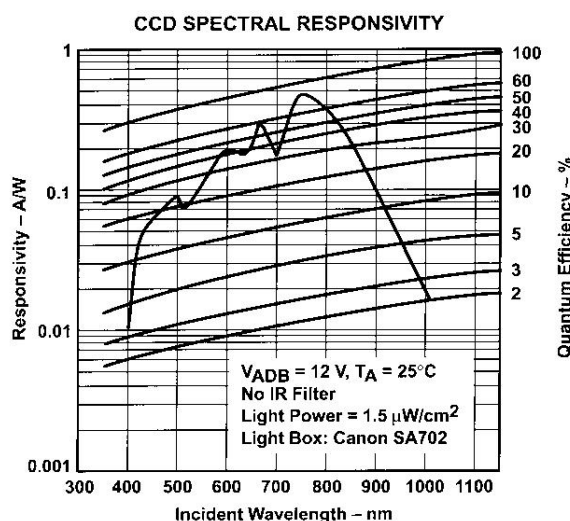


FIGURE 2.12. CCD spectral responsivity.

therefore a random variable. The standard deviation of this process is the photon noise and needs to be considered when the scenes are relatively dim and few photons strike a pixel. This noise tends to follow Poisson statistics, but is often masked by other noise sources.

2.7. Dark current and dark current noise. There is normally leakage current present in each CCD cell that alters the charge acquired during a cycle. Dark current is dependent on the CCD make up, defects, temperature and operating voltages. Dark current noise is the statistical variation in dark current. The total charge introduced by dark current increases with charge integration time while the relative noise tends to decrease.

2.8. CCD defects. There are a variety of possible CCD defects. These are normally graded as point or cluster defects, column or row defects, charge traps or hot pixels. Such defects can occur at various locations within the photosensitive, storage or shift register regions of the CCD and may result in varying degrees of signal degradation.

2.9. Clocking noise. A small amount of charge is injected into each CCD cell when it is cycled, dependent somewhat on the clocking sequence, frequency and waveform. Normally, the sharper the clock waveform, the larger the signal disturbance. Special attention during design is required to minimize this noise factor.

2.10. Reset noise. CCD readout typically involves the charging of a capacitance through a switch (reset) device. From thermodynamic theory, the voltage on the floating diffusion capacitor is not the same after every reset, even if there is no charge collected from the previous detection. The noise in this reset process due to thermal noise is directly reflected in the output signal. The reset level is a random value whose standard deviation is the reset noise.

2.11. Charge transfer efficiency. Charge transfer efficiency is a measure of how well the charge in a particular pixel is advanced to the next cell during a cycle. This efficiency is dependent on the design and quality of the CCD as well as the operating voltages and frequencies.

Smearing effects can arise when the stored charge packets are shifted through the imaging region of the CCD towards the output node. During this readout time, the photo-generation of electrons continues. This can contaminate charge packets that have originated from one part of the image with additional electrons from another. To minimize the smear effect, it is crucial to ensure that the shift register is read out in a shorter time interval than that used during the collection of signal electrons (the integration time). Hence, for a given integration time, low smear requires a high readout rate.

Another source of smear arises when electrons generated in the silicon are “captured” by a nearby storage gate rather than the storage gate upon which the light was incident. This is a statistical effect due to the random process of diffusion by the electrons towards the storage region. As longer wavelengths are absorbed more deeply in the sensor the smear tends to worsen.

2.12. CCD readout and charge conversion limitations. There are a number of effects that limit the quality of the charge measurement process. As the charge is collected, transferred along the chain of CCD cells and shifted out through the charge-to-voltage converter, various distortions are introduced.

2.13. Charge-to-voltage conversion limitations. The charge packets in the CCD register are too small to be transferred directly to a standard video signal without significantly degrading the signal-to-noise ratio. Instead, a small sensing capacitor is used to convert the signal charge to a voltage that can be amplified before transmission. The

performance of this stage is judged in terms of its charge conversion efficiency. The conversion of the charge packets into a voltage signal requires several stages of analog signal processing. These stages are subject to a range of performance limitations including noise, bias shifts, non-linear amplitude and frequency responses, timing artifacts, power supply effects and temperature drift. Even with careful design, these issues cannot be completely eliminated.

2.14. Video encoding. The output produced by the “charge-to-voltage” conversion is then encoded into an analog video signal. The encoding process adds vertical and horizontal timing components according to a standard video format. Downstream system components utilize this timing information to interpret which portions of the signal correspond to which regions in the original image.

The video encoding process produces another level of image distortions. These distortions cause variations in the signal timing, amplitude, frequency response and noise. Some of these may be spatially invariant such that they are included in the camera lens calibration, or they may fluctuate with the operating conditions and time. The net result is that there is degradation in the ability to determine the position and relative intensity of image features.

2.15. General CCD aberration contributions. Under most operating conditions, unless proven otherwise, CCD aberrations can be ignored for the camera calibration exercise since both internally generated noise and quantization error are below the tolerance of standard TV resolution. An exception might be the uncertainty in the horizontal scale factor discussed by Lenz and Tsai in [35] for video frame grabbing.

3. CCD and lens contributions

Figure 2.13 illustrates the theory underlined in Figure 2.1, reinforcing the idea that under normal conditions the CCD together with the lens aberrations, only produce a localized systematic filtered copy of the original signal after the distortion.

The luminance profile of Figure 2.13a is a section of what an edge image intensity profile would look like if an ideal lens were used instead of the real one. Due to the limited aperture of the lens, the actual luminance profile is a filtered version of the ideal one, as

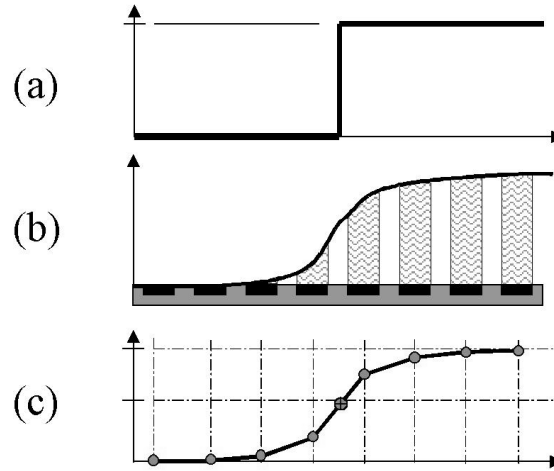


FIGURE 2.13. Subpixel edge detection based on linear interpolation. (From [43])

shown in Figure 2.13b. When the light reaches the array of the photosensors of the CCD camera the image is spatially sampled. The luminance samples actually collected by the CCD array depend on the light that falls on the whole photosensitive area, as shown in the bottom curve in 2.13c. At this stage, different image processing techniques can be used to estimate the subpixel location of the ideal edge.

The camera calibration operation, placed in the context of the image formation system just described, consists of using extracted image feature locations subject to shifting, blurring and additive noise to unveil fundamental information about the camera. From the sections above, it is implied that only the nonlinear shifting represented by the first block in Figure 2.1, should be compensated for. Therefore, all other effects described should only be taken into account during the image processing stages where the feature locations are extracted.

4. Conclusion

This chapter discussed the fundamentals of the CCD image formation theory. The concepts presented are helpful to understand and assess the different strategies used to calibrate CCD cameras.

The following chapter sets the basis for the mathematical equations describing the common camera parameters that will be used for the remainder of this thesis.

CHAPTER 3

Mathematical concepts for camera calibration

This section describes the general pinhole camera model, including the intrinsic and extrinsic calibration parameters.

1. The camera model and coordinate systems for the calibration

Figure 3.1 illustrates the basic geometry of the pinhole camera model adopted for this study.

X_T , Y_T , and Z_T are the axes of the right-handed 3D Target Coordinate System (TCS) fixed with respect to the payload or the calibration array. Units used are inches. Target locations are defined in the TCS. The origin of the TCS is the geometric center of the set of target locations. The axes are chosen to be approximately parallel to the Camera Coordinate System (defined in the next paragraph) in order to ensure convergence of the photogrammetric minimization algorithm used for this study. Specific TCSs used for this study are defined as needed in later chapters.

The axes X_C , Y_C , and Z_C define the right-handed 3D Camera Coordinate System (CCS), which is centered at the camera projection center (defined later), with the X_C axis being the same as the camera optical axis. The Y_C axis is parallel to the horizontal axis of the image (y), from left to right, and the X_C axis is parallel to the vertical axis (z), from top to bottom. Measurements in the CCS are expressed in inches.

The creation of another 3D coordinate system attached on a fixed element of the camera housing is necessary to define the location and orientation of the camera optical center. This additional vector (X_H , Y_H , Z_H) is the right-handed 3D Housing Coordinate System

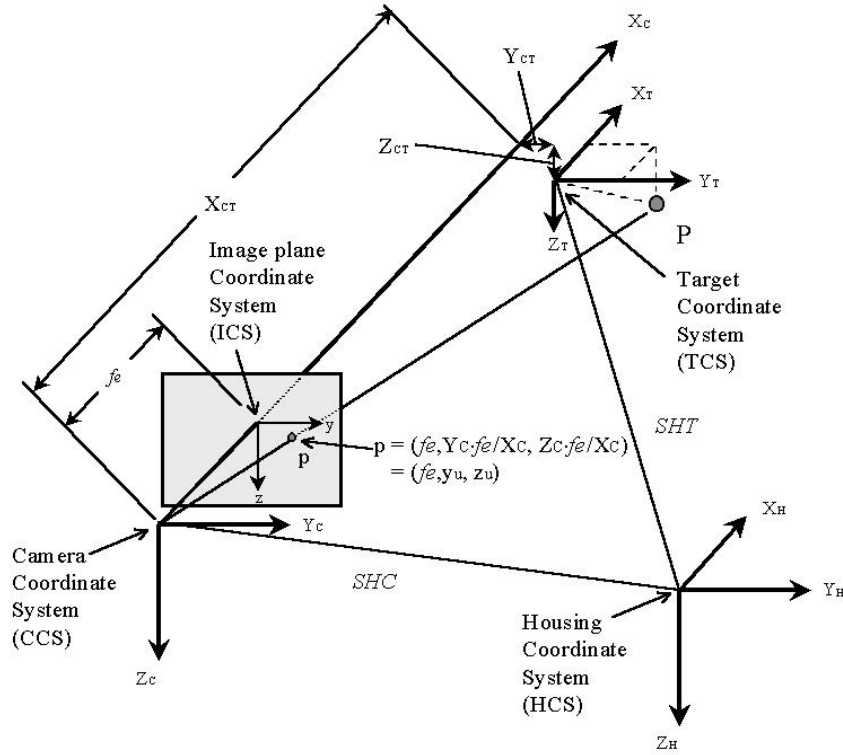


FIGURE 3.1. Pinhole camera geometry.

(HCS) and is also expressed in inches. The origin of the HCS is at a fixed and physically measurable point on the camera lens assembly. For this thesis, the origin is taken as the center of the circle formed by the front surface of the lens assembly (not the glass of the lens). The axes of the HCS are typically set to be parallel to fixed and measurable edges on the camera housing. To use the HCS in an operational context, the relationships with other coordinate systems related to the physical camera (e.g. camera pan/tilt axis) are obtained from engineering drawings.

The distance f_e , the camera effective focal length, is the distance between the frontal image plane and the origin of the camera coordinate system. The term “effective focal length” is used because the distance from the projection center to the image plane is defined in the pinhole camera model and does not necessarily have the same value as the focal length measured optically (see Chapter 2). Dimensions for the effective focal length and the target image locations are reported in millimeters.

The Frame Coordinate System (FCS), illustrated later in Figure 3.3, and expressed by y_F and z_F , is the scan-ordered measurement coordinate system for the sampled video image. The origin of the FCS is a point above the upper left corner of the image, at a location dependent on the hardware design. The y_F axis is in the direction of the video raster scan, and the z_F axis is perpendicular to y_F and positive down.

The axes y and z represent the 2D coordinate system within the image plane and define the Image plane Coordinate System (ICS). The ICS origin is located at the intersection of the X_C axis and the image plane. For CCD cameras the y and z axes are parallel to the y_F and z_F axes of the FCS defined above. Units in the ICS are millimeters. Since the units for the FCS are expressed in pixels and lines for the discrete image in the frame memory, additional parameters need to be specified (or calibrated) which relate the image coordinate in the FCS to the ICS. The vector (y_U, z_U) is the image coordinate p in the ICS of an object point P if a perfect pinhole camera model is used.

In Figure 3.1, the position and orientation of the CCS with respect to the HCS is called the SHC vector. Similarly, the position and orientation of the TCS with respect to the HCS is called the SHT vector. The SHC vector also represents the transformation from the HCS to the CCS and the SHT vector represents the transformation from the HCS to the TCS.

The next four subsections detail the mathematical terms underlying the whole perspective projection model.

1.1. Rigid body transformation from the TCS to the CCS. The unique transformation expressing the transition of an object point P_i from the target coordinate system to the camera coordinate system is given by:

$$\begin{bmatrix} X_C(i) \\ Y_C(i) \\ Z_C(i) \end{bmatrix} = R \cdot \begin{bmatrix} X_T(i) \\ Y_T(i) \\ Z_T(i) \end{bmatrix} + T.$$

where R is a 3 by 3 rotation matrix and T is the translation vector expressed by:

$$R = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} X_{CT} \\ Y_{CT} \\ Z_{CT} \end{bmatrix}.$$

In the above equation, the translation vector formed with X_{CT} , Y_{CT} and Z_{CT} expresses the location of the target coordinate system with respect to the reference of axes centered at the projection center. The terms α_{ij} form the direction cosines defining the angles between the target coordinate system and the camera coordinate system. The six degree-of-freedom transformation obtained with R and T is called the SCT vector.

1.1.1. *Rotation matrix and Euler angles.* As we just noted, the positional parameters are X_{CT} , Y_{CT} and Z_{CT} . For orientation, a choice of three independent direction cosines has to be made. The independent triplet is chosen to be α_{12} , α_{13} , and α_{23} . The remaining direction cosine parameters are given by the following relations given in [64] and calculated in the sequence shown:

$$\begin{aligned} \alpha_{11} &= \sqrt{1 - \alpha_{12}^2 - \alpha_{13}^2}, \\ \alpha_{33} &= \sqrt{1 - \alpha_{13}^2 - \alpha_{23}^2}, \\ \alpha_{21} &= \frac{-(\alpha_{23}\alpha_{13} + \alpha_{11}\alpha_{23}\alpha_{13})}{1 - \alpha_{13}^2}, \\ \alpha_{22} &= \frac{\alpha_{33}\alpha_{11} - \alpha_{13}\alpha_{12}\alpha_{23}}{1 - \alpha_{13}^2}, \\ \alpha_{31} &= \alpha_{12}\alpha_{23} - \alpha_{22}\alpha_{13}, \\ \alpha_{32} &= \alpha_{13}\alpha_{21} - \alpha_{23}\alpha_{11}. \end{aligned}$$

More intuitive orientation measurements can be extracted from the previous matrix, which define the rotations (in degrees) about the three axes forming the coordinate system. Following the example from [16, page 42], the Pitch-Yaw-Roll (θ, ψ, ϕ) ¹ Euler rotation angles would produce the following rotation matrix R :

¹Rotation sequence used by NASA for RMS operations.

$$R = \begin{bmatrix} \cos \theta \cos \psi & -\cos \phi \cos \theta \sin \psi + \sin \theta \cos \phi & -\sin \phi \cos \theta \sin \psi + \cos \phi \sin \theta \\ \sin \psi & \cos \phi \cos \psi & \sin \phi \cos \psi \\ \sin \theta \cos \psi & \cos \phi \sin \theta \sin \psi - \sin \phi \cos \theta & \sin \phi \sin \theta \sin \psi + \cos \phi \cos \theta \end{bmatrix}.$$

The solution for extracting the Pitch-Yaw-Roll Euler angles using the Atan2 function² follows:

$$\begin{aligned} \psi &= \text{Atan2}(\alpha_{21}, \sqrt{\alpha_{23}^2 - \alpha_{22}^2}), \\ \theta &= \text{Atan2}\left(\frac{-\alpha_{31}}{\cos \psi}, \frac{\alpha_{11}}{\cos \psi}\right), \\ \phi &= \text{Atan2}\left(\frac{-\alpha_{23}}{\cos \psi}, \frac{\alpha_{22}}{\cos \psi}\right). \end{aligned}$$

If $\psi = 90^\circ$, then

$$\begin{aligned} \psi &= 90^\circ, \\ \theta &= \text{Atan2}(\alpha_{32}, -\alpha_{12}), \\ \phi &= 0.0^\circ. \end{aligned}$$

Also, if $\psi = -90^\circ$, then

$$\begin{aligned} \psi &= -90^\circ, \\ \theta &= \text{Atan2}(-\alpha_{32}, \alpha_{12}), \\ \phi &= 0.0^\circ. \end{aligned}$$

Note that the rotation parameters (pitch (θ), yaw (ψ), and roll (ϕ)) can be solved directly without using the direction cosines. This approach is currently used in the Space Vision System (SVS), although a yaw-pitch-roll unpacking sequence is employed.

²The definition of the Atan2 function was borrowed from [16, page 41].

1.2. Transformation from the CCS to ideal undistorted ICS. When projected on the image plane, the point P_i becomes p_i and is represented by:

$$p_i = \left(f_e, \quad y_U(i), \quad z_U(i) \right)^T = \left(f_e, \quad \frac{f_e Y_C(i)}{X_C(i)}, \quad \frac{f_e Z_C(i)}{X_C(i)} \right)^T.$$

As a result, the previous vector and the vector from Section 1.1 are related by a scalar quantity $k(i)$, representing the localized image scale and defined by:

$$k_i = \sqrt{\frac{X_{CT}^2(i) + Y_{CT}^2(i) + Z_{CT}^2(i)}{f_e^2(i) + y_U^2(i) + z_U^2(i)}}.$$

The set of previous equations is now given by:

$$(3.1) \quad k_i \cdot \begin{bmatrix} f_e \\ Y_U(i) \\ Z_U(i) \end{bmatrix} = \begin{bmatrix} X_{CT}(i) \\ Y_{CT}(i) \\ Z_{CT}(i) \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \cdot \begin{bmatrix} X_T(i) \\ Y_T(i) \\ Z_T(i) \end{bmatrix}.$$

Solving for k_i using the system of Equations 3.1 and substituting back into the second and third, produces the colinearity relations of photogrammetry for the pinhole camera model:

$$\begin{aligned} \frac{y_U(i)}{f_e} &= \frac{Y_{CT}(i) + Y_T(i)\alpha_{21} + Y_T(i)\alpha_{22} + Z_T(i)\alpha_{23}}{X_{CT}(i) + Y_T(i)\alpha_{11} + Y_T(i)\alpha_{12} + Z_T(i)\alpha_{13}}, \\ \frac{z_U(i)}{f_e} &= \frac{X_{CT}(i) + Y_T(i)\alpha_{31} + Y_T(i)\alpha_{32} + Z_T(i)\alpha_{33}}{X_{CT}(i) + Y_T(i)\alpha_{11} + Y_T(i)\alpha_{12} + Z_T(i)\alpha_{13}}. \end{aligned}$$

The colinear conditions contain a total of six unknowns - three defining the position, as given by X_{CT} , Y_{CT} , Z_{CT} , and three selected independent direction cosine terms (α_{12} , α_{13} , and α_{23}) to define the orientation.

1.3. Lens distortion correction. Figure 3.2 is a representation of the distortions introduced by a lens. The camera is first aimed at the calibration board designed with evenly spaced circular target elements. An image of the board, acquired with the camera and a video digitizer, shows some discrepancy between the measured dot locations and the

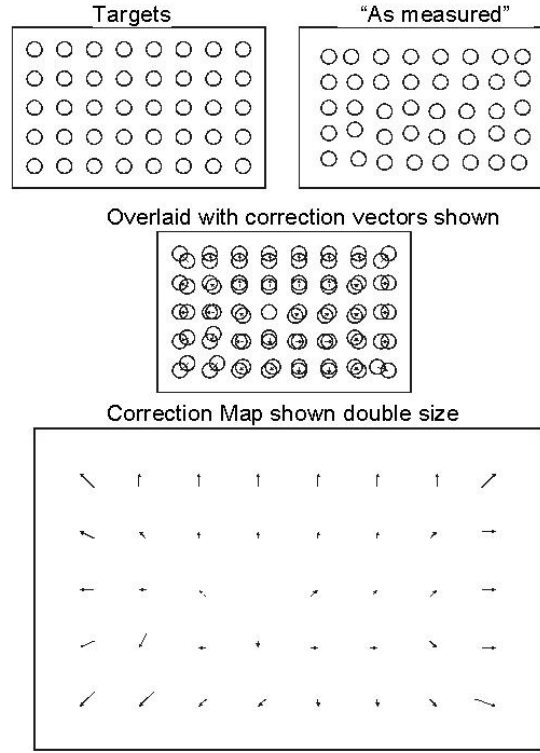


FIGURE 3.2. Centroid biases caused by lens distortions.

locations predicted from the perspective projection model (pinhole camera). The *distortion correction map* defines the displacement from the distorted to the ideal dot locations.

The vector (y_D, z_D) is the actual measured image coordinate, which differs from (y_U, z_U) , the undistorted centroid, due to lens distortions. Hence, the distortion-free centroids are recovered by adding an extra correction value (pixels or lines) to the measured distorted ones. D_y and D_z are generic correction terms derived from a modeling of the lens distortions. D_y and D_z are functions of the measured centroids (y_D, z_D) . Different methods to obtain these terms can be found later in this document. The transformation formulas are:

$$y_U(i) = y_D(i) + D_y,$$

$$z_U(i) = z_D(i) + D_z.$$

1.4. Transformation from the FCS to the ICS. The target centroid, as returned by the scanning and digitization hardware, is expressed in pixels and lines having its natural

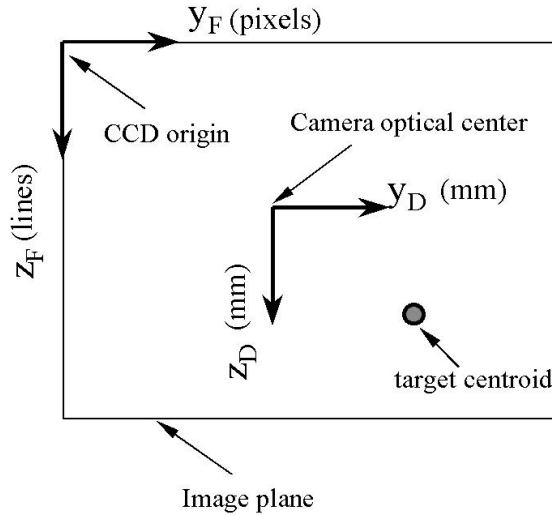


FIGURE 3.3. Centroid location based on two coordinate systems in the camera image plane.

origin at the top left corner of the image plane. This centroid representation differs from what is desired in the photogrammetry colinear equations where the target image locations have to be expressed in millimeters (same units as f_e) from the camera optical center. Figure 3.3 illustrates this concept.

The transformation from the FCS to the ICS is given by:

$$\begin{aligned} y_D(i) &= (y_F(i) - C_y) \cdot G_y, \\ z_D(i) &= (z_F(i) - C_z) \cdot G_z. \end{aligned}$$

In the previous equations (C_y , C_z) fix the origin of the ICS in pixels and lines in the FCS. As defined in [68, 67], as many as fifteen different types of image centers can be considered: center of digitizer, center of sensor, center of the radial distortion, center of perspective projection, center of field of view, center of focus, etc. In the context of this thesis, for the methods implemented, the optical center (also called principal point) is defined as the intersection of the projection axis of the camera with the image plane. In the pinhole camera model, the optical center is the intersection of the line through the projection center and perpendicular to the image plane defined by (y_U, z_U) . Therefore, the location of the optical center, in object space with respect to the CCS is:

$$\begin{pmatrix} X_C & Y_C & Z_C \end{pmatrix}^T = \begin{pmatrix} f_e & 0 & 0 \end{pmatrix}^T.$$

It is important to note that the physical image plane (CCD array) is not necessarily perpendicular to the projection axis defined by the pinhole camera.

The terms G_y and G_z are called image gain ratios (expressed in mm/pixel and mm/line respectively) and establish the conversion factor between measurements in pixels (or lines) and the unit of the effective focal length (expressed in mm). Typically, substitute values for the calibration of the gain ratios of any CCD imager can be derived from the technical specification sheet under the terms cell size, or sampling cell size. However, as described in [55], it is necessary to include a non-dimensional term S_y for the uncertainty inherent to the design of the hardware used for scanning and digitizing the lines in the composite video signal. This is due to a variety of factors, such as slight hardware timing mismatch between image acquisition hardware and camera scanning hardware, or the imprecision of the timing of the TV scanning itself. Even a one-percent difference can cause three to five pixel errors for a full resolution frame. Using the uncertainty term, the previous equations from the dot locations become:

$$\begin{aligned} y_D(i) &= (y_F(i) - C_y) \cdot G_y \cdot S_y, \\ z_D(i) &= (z_F(i) - C_z) \cdot G_z. \end{aligned}$$

2. Extrinsic and intrinsic camera parameters

The task behind the camera calibration is to compute the camera's intrinsic and extrinsic parameters based on a number of points whose locations in the TCS are known, and whose image coordinates are measured. Although a calibration object may need to be repositioned several times during the calibration to recreate a 3-D reference system, only one location is chosen as the "reference" calibration position, and is designated by T_R . Considerations for choosing T_R are presented in a subsequent section. A description of the parameters constituting the two classes is provided in the following two subsections:

2.1. The extrinsic camera parameters. The six parameters from Section 1.1 used to define the SCT_R vector constitute the extrinsic parameters. SCT_R is determined at the reference location T_R . The six parameters X_{CT_R} , Y_{CT_R} , Z_{CT_R} and associated α_{12_R} , α_{13_R} and α_{23_R} of the SCT_R vector form the extrinsic parameters for this study.

For operational use, the SHC is an important vector that one needs to know to measure the physical camera body aiming. It sets the constant transformation between the orientation of the camera pan-tilt unit and the CCS. During the last phase of the calibration, by physically measuring the SHT_R vector and by using the calibrated SCT_R vector, the location and orientation parameters can be retrieved to estimate the SHC vector relating the CCS to a physical static point on the camera. The transformation cycle is given by:

$$SHC = SHT_R \cdot (SCT_R)^{-1}.$$

For this study, due to limited laboratory resources, only two of the SHT_R six degrees-of-freedom was measured. Specifically, using a tape measure and an inclinometer, $X - SHT_R$ and $Roll - SHT_R$ were obtained. The other degrees-of-freedom were assumed to be zero.

A more efficient and precise method to calculate SHC vectors is presented in [59]. The approach was designed to work with camera Pan-Tilt Units (PTU) equipped with attitude indicators. Starting with the knowledge of a short distance between a fixed point on the camera and the rotation center of the PTU, the camera undergoes a series of small rotations keeping a static target array in its field-of-view. After each transformations, the SCT vectors between the camera and the target array are recorded. Using the gathered data from at least three measurement points, all the necessary information is available to retrieve the SHC vector using a non-linear minimization approach. Equation 3.2 contains the mathematical expression used to uncover $SHC_{1,2}$, which represents the transformation of the CCS (SC_1C_2) produced by a known rotation (R) of the PTU. The equation is:

$$(3.2) \quad SC_1C_2 = (SHC_{1,2})^{-1} \cdot R \cdot SHC_{1,2}.$$

2.2. The intrinsic camera parameters. The parameters used in Sections 1.2, 1.3, and 1.4 for the transformation from the 3D object coordinates in the FCS to the CCS are

called intrinsic parameters. At a minimum, the group includes: the effective focal length, the lens distortion factors, the image plane gain ratios and the image optical center.

As seen above for the colinear photogrammetry equations, the units for the centroids and the effective focal length are chosen to be millimeters. Two different approaches can be envisioned to determine the parameters describing the size of the image plane and the effective focal length. The first approach is to assume that the focal length is known (constant value) and to determine the image gain ratios during the calibration. The second route is to assume the gain ratios to be constant and to solve for the unknown effective focal length and TV scanning uncertainty factor S_y . These two approaches were tested for this thesis.

3. Conclusion

This chapter defined the different parameters necessary to define the complete camera characteristics regardless of the calibration method employed. The next chapter describes the calibration equipment and some procedural requirements necessary to maximize the chances of performing meaningful camera calibrations.

CHAPTER 4

Camera calibration requirements

In this chapter it is argued that the accuracy obtained during the camera calibration process not only depends on the calibration technique, but is also influenced by environmental factors. These factors can be grouped into three categories: camera calibration equipment, camera setup, and laboratory conditions.

Please note that the impact analysis of various error sources on camera calibration accuracy is not addressed. The recommendations presented are solely based on experience and theoretical analyses.

A list of parameters to be monitored during the calibrations is included. The verification of those indices adds confidence and legitimacy to the calibration results.

The last section of this chapter contains an overview of the test equipment and video processing software/hardware used for this study.

1. Camera calibration equipment

The calibration rig is the fundamental component of the calibration task. The level of tolerance and functionality specified to build the apparatus has a direct effect on the accuracy of the camera calibration task. It serves as “ground truth” for establishing the projective scaling obtained with a given camera.

Three main constituents form the calibration rig. First, the *Camera Mounting Apparatus* (CMA) is used to securely support and control the position and orientation of the camera. Second, the *Calibration Object* (CO) defines the physical standard against which the ideal target locations are measured. It is shaped such that a maximum of calibration

features can occupy the field-of-view of a standard 4:3 aspect ratio camera. Lastly, artificial lighting systems are used to evenly illuminate all features of the CO.

A multitude of design variants exist between calibration rigs but fundamentally the functionality of the different components remains the same for all setups. The selection and design generally depends on the camera calibration method chosen, but is also influenced by other factors such as simplicity, cost, desired accuracy, speed, and autonomy.

1.1. Construction tolerances. Ideally, the size of the calibration object should be similar to one used during the operation. However, this is not always practical, nor feasible, and therefore small scale calibration objects have to be used instead. In such cases, construction inaccuracies above a level detectable by the sensor, are directly amplified and can sometimes cause accuracy problems for the true scale operation. The simplest solution to this problem is to design a calibration object with sufficient tolerances, made out of stable and rigid materials.

The allowable magnitude in feature location error, such that it cannot be detected, is proportional to the camera resolution and to the tangent of the angle between the feature line-of-sight and the camera optical axis. Hence, inaccuracy has a greater influence on the outer regions of the field-of-view and for shorter focal length lenses. Figure 4.1, together with Equations 4.1 and 4.2, illustrate a practical and straightforward methodology to define the required accuracy for any calibration object.

The box surrounding the calibration feature in Figure 4.1 defines the allowable area representing the manufacturing tolerances. The calculation is performed using the width of the narrowest side of the camera pixel. Referring to the nomenclature of Figure 4.1, four unknowns, a , b , c , and d , need to be determined. The final tolerance on any one object feature location can be specified using: X_{-c}^{+a} , Y_{-d}^{+b} and Z_{-d}^{+b} . A single tolerance applicable for all elements is established from the one element which is the closest in X and the furthest in Y and Z . Tolerance specifications are given by: X_{-c}^{+a} , Y_{-b}^{+b} and Z_{-b}^{+b} . Equations 4.1 and 4.2 present, step by step, the process for recovering those parameters. The coordinate system follows the one adopted for Chapter 3. The equations are:

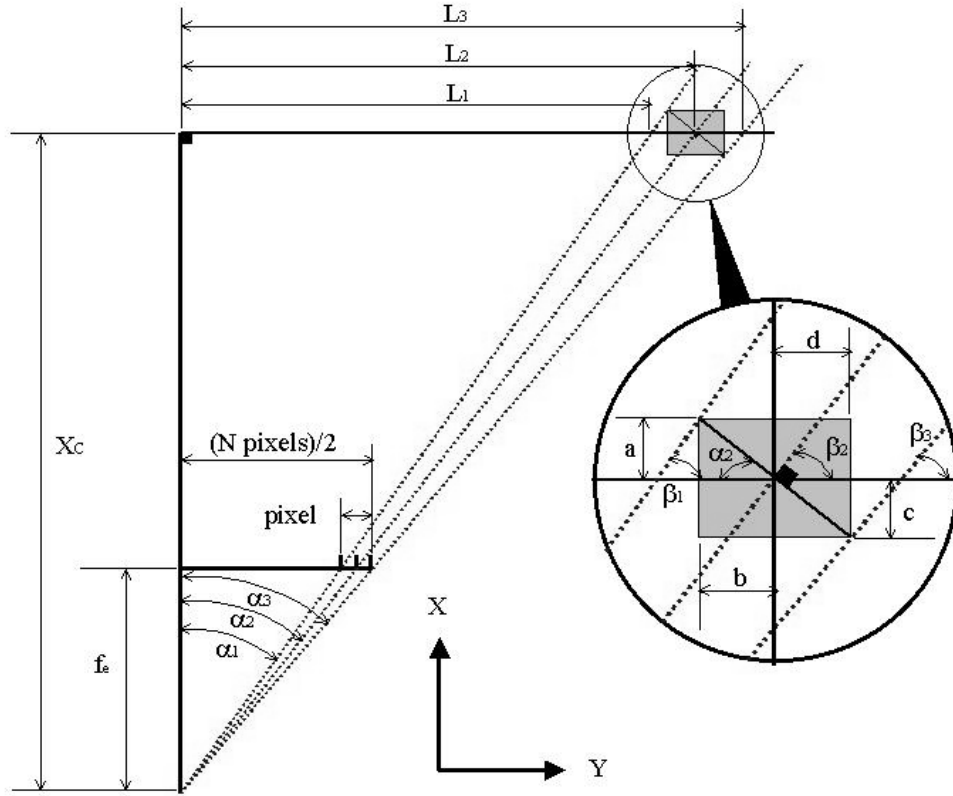


FIGURE 4.1. Graphical representation of the calibration object construction tolerances.

$$\tan \alpha_1 = \frac{[(N_{pixels}/2) - 0.5] \cdot G_y}{f_e}, \quad \tan \alpha_2 = \frac{[(N_{pixels}/2)] \cdot G_y}{f_e}, \quad \tan \alpha_3 = \frac{[(N_{pixels}/2) + 0.5] \cdot G_y}{f_e},$$

$$(4.1) \quad L_1 = \tan \alpha_1 \cdot X_C, \quad L_2 = \tan \alpha_2 \cdot X_C, \quad L_3 = \tan \alpha_3 \cdot X_C$$

$$\tan \beta_1 = \frac{\pi}{2} - \alpha_1, \quad \tan \beta_3 = \frac{\pi}{2} - \alpha_3.$$

After simplifications, the values for a , b , c , and d in the order extracted, are given by:

$$(4.2) \quad \begin{aligned} b &= \frac{(L_2 - L_1) \cdot \tan \beta_1}{(\tan \alpha_2 + \tan \beta_1)}, \\ a &= \tan \alpha_2 \cdot b, \\ d &= \frac{(L_3 - L_2) \cdot \tan \beta_3}{(\tan \alpha_2 + \tan \beta_3)}, \\ c &= \tan \alpha_2 \cdot d. \end{aligned}$$

The tolerance values just mentioned are “absolute”, and therefore one should consider possible mechanical play or other error sources which could be introduced by moving parts.

1.2. Calibration array pattern. As shown, the construction accuracy of the calibration object is crucial for camera calibrations. Also important is the optical nature of the landmarks on the object, which must be selected to create a high signal-to-noise ratio in the video signal for non-ambiguous feature detection. The calibration objects are generally made with one, or a combination of, planar surfaces with high contrast patterns which are painted, printed or glued on. The choice of calibration pattern depends on the type of image processing used to measure image feature centroids.

Popular centroid measurement methods for camera calibrations include *blob detection* and *edge detection*. These are now presented.

1.2.1. Blob detection. This technique refers to the detection, in greyscale images, of fully confined regions of similar video intensities. Using a particular video thresholding algorithm, a binary video is generated which circumscribes the target foreground from the image background. Blob detection algorithms need a sufficient target background width to accurately discretize a blob. As shown in Figure 4.2, a minimum of three background pixels between each target is generally sufficient, given standard video noise levels.

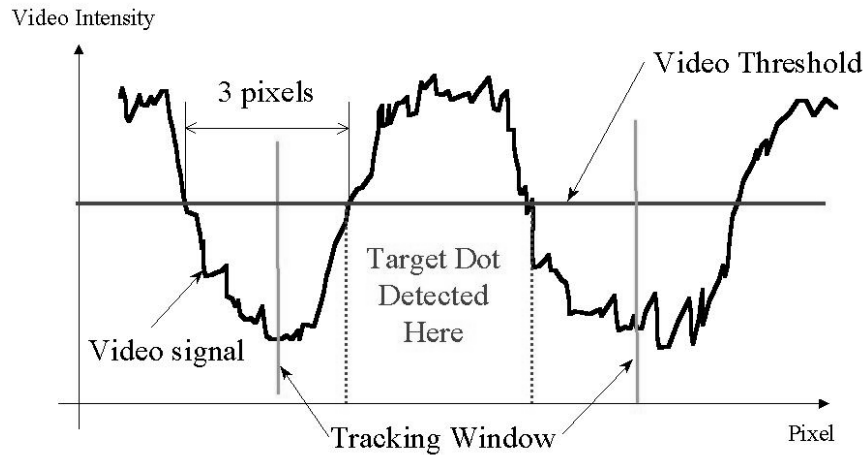


FIGURE 4.2. Video thresholding for blob detection.

The centroid of the target is measured with respect to the general image coordinate system using a “center of mass” calculation technique applied on the binarized video lines

forming a blob. The equations in 4.3 contain a simple methodology to measure a blob centroid (\bar{y}, \bar{z}) from binarized, raster scanned, video lines. The necessary information to carry out the calculation is the number of horizontal line segments forming the blob (N_{seg}) and each segment's video line number (Lid_i), together with the beginning and ending columns ($Cst_i, Cend_i$). The equations are:

$$\begin{aligned}
 (4.3) \quad Area &= \sum_{i=1}^{N_{seg}} (Cst_i - Cend_i), \\
 \bar{y} &= \frac{1}{Area} \sum_{i=1}^{N_{seg}} (Lid_i \cdot (Cst_i - Cend_i)), \\
 \bar{z} &= \frac{1}{Area} \sum_{i=1}^{N_{seg}} \left(\frac{(Cst_i - Cend_i)^2}{2} \right).
 \end{aligned}$$

Naturally, the target shape which provides the most centroid robustness against CCD aliasing and noise is the circle. As explained in [31], the edge curvature of a circle tends to average out the quantization effects of the CCD in proportion to its diameter. The formula for estimating the centroid location accuracy is presented later in this section.

In spite of this, increasing the target size introduces complications. Hence, a large target reduces the calibration range on the image plane since the target more readily intersects the border of the field-of-view, and also the computation of the centroid is sensitive to the viewing angle. In addition, since discrete calibration points are used, the calibration technique must employ an interpolation scheme between those reference elements to characterize the entire camera field-of-view. Since the interpolation generally becomes less accurate as the distance between the reference points is increased, the accuracy of the method is further improved as the grid density is increased, which is at the expense of reducing the target sizes. A compromise has to be established between an optimal target size for the system to measure accurate centroids and a maximum number of the calibration circles in the calibration array to minimize interpolation errors.

Figure 4.3 illustrates three cases of first order radial lens distortion where a standard bi-linear interpolation over a square grid is used. The centroid shift is measured for a target located equi-distant between four reference points. The magnitude is given as a

function of the distance between those four reference points on a flat calibration array. Naturally, for radial distortions, the interpolation error increases when the distance from the distortion center increases. For this simulation, we measured the interpolation error for a target element located at the bottom right corner of the array (zone of maximum image distortion), precisely at coordinate ($Y_C = 11.375$ inches, $Z_C = 7.875$ inches) at a distance from the camera (X_C) which depends on the camera focal length and would allow all five dots to remain in the field-of-view. Further details about this test setup can be found in Table 6.1. The data presented here are derived from three test cases studied for this thesis.

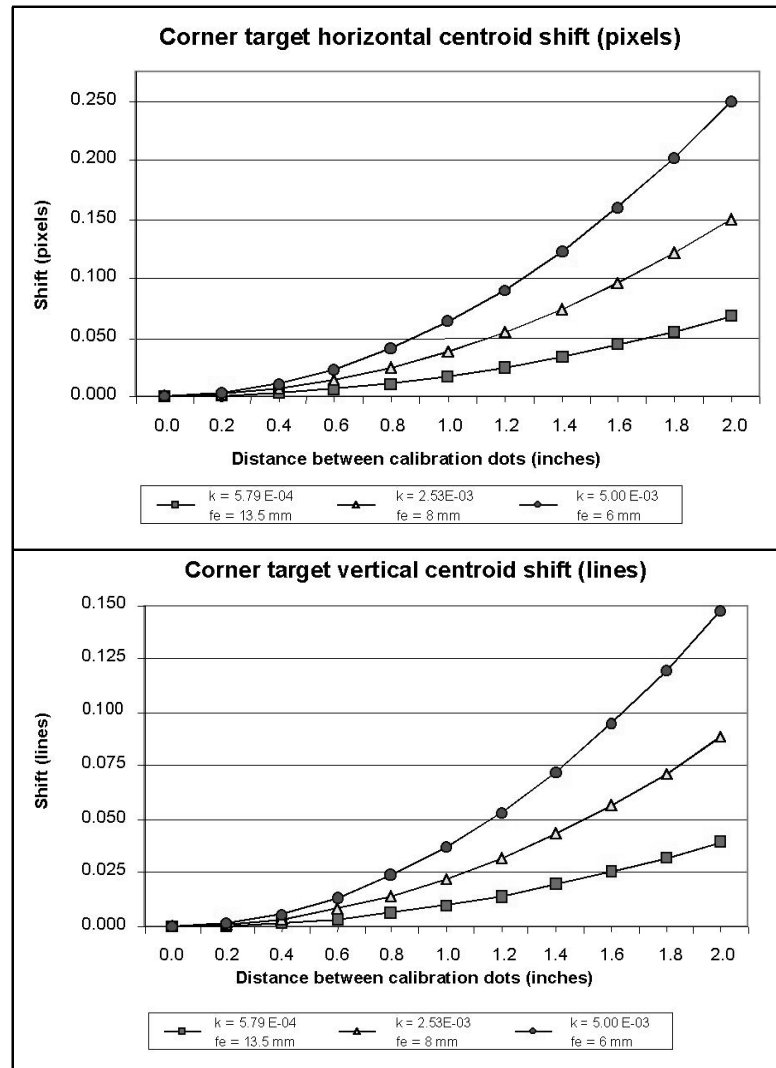


FIGURE 4.3. Centroid shift due to bi-linear interpolation.

Figure 4.4 illustrates three cases of first order radial lens distortion where errors are caused by local distortion affecting the shape of the projected circular object. As in the previous case, the target element is located at the bottom right corner of the calibration array, and now has the coordinate ($Y_C = 12.25$ inches, $Z_C = 10.50$ inches) at a distance from the camera (X_C) which depends on the camera focal length and would allow the dot to remain at the very edge of the field-of-view. As before, details can be found in Table 6.1.

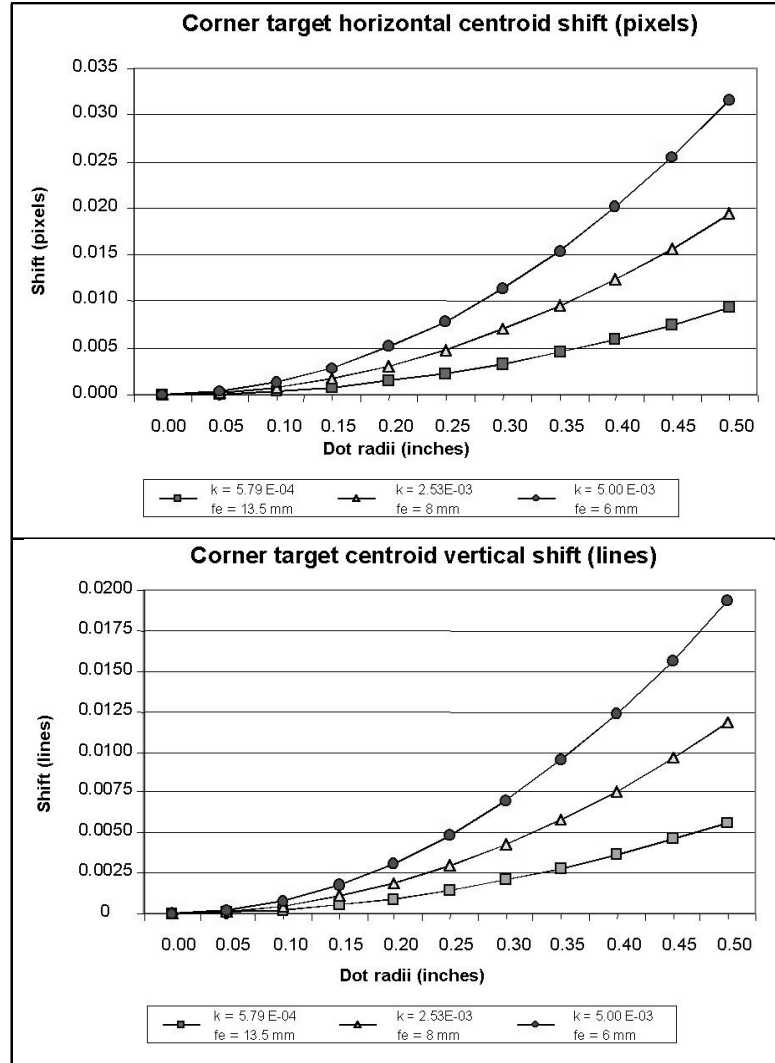


FIGURE 4.4. Centroid shift versus target radius.

It is worth mentioning that the centroid shift errors presented in Figure 4.4 will approach zero if the operational dot sizes are the same as the calibration dot sizes.

Last, but not least, is the formula given in [31] for estimating the centroid location accuracy given the target diameters d_y and d_z in pixels and lines:

$$(4.4) \quad \begin{aligned} \delta_y &= \frac{0.222}{\sqrt{d_y}}, \\ \delta_z &= \frac{0.222}{\sqrt{d_z}}. \end{aligned}$$

A visual representation of equations 4.4 is given in 4.3 and 4.4. In order to be compatible with previous figures, the target size was transformed from pixels and lines to inches.

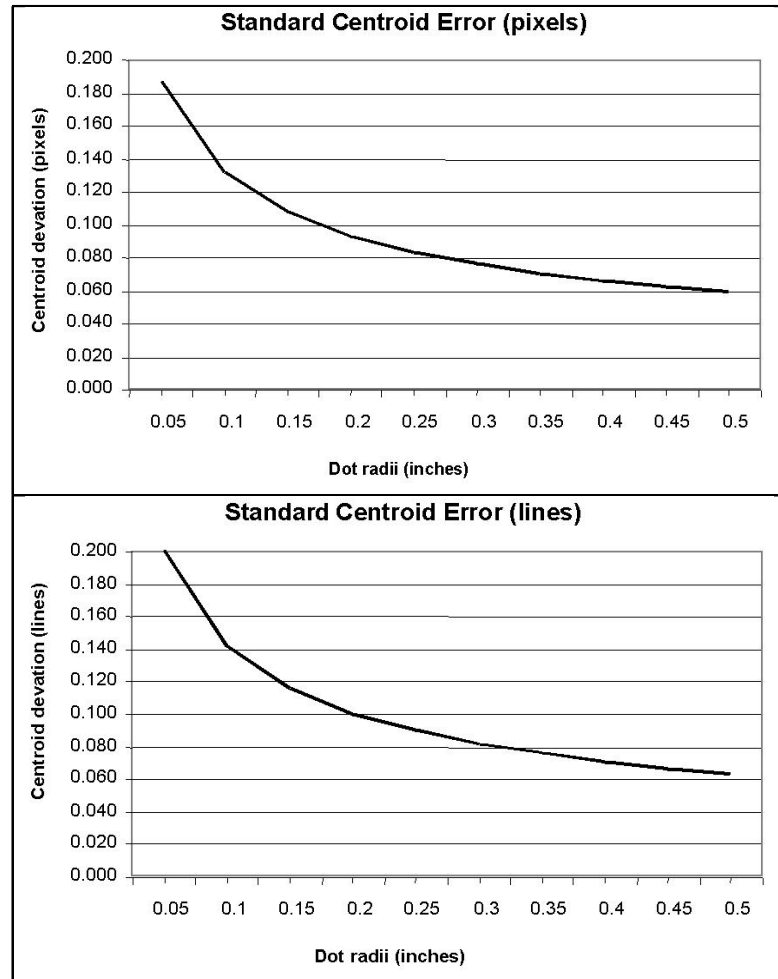


FIGURE 4.5. Standard centroid measurement error from [31].

Finally, an acceptable compromise to select the size and distance between each target, is to assume a distance, center-to-center, equal to three times the dot radius and, for a given focal length and distortion, look for the intersection point that minimizes the errors using Figures 4.3 and 4.5. For example, with the 6 mm lens used for this study, the ideal target diameter would be approximately 0.7 inches.

1.2.2. Edge detection. Edge detection is the art of accurately positioning a step-like intensity transition in a video signal resulting from sharp transitions in the scene being imaged. For camera calibration, edge detection and morphological operators are powerful tools used to distinguish the intersection of line segments, line mid-points, or sample points on the perimeter of a well defined primitive. Such landmarks reveal purposely defined features of the calibration object which are used as reference points during the calibration.

Normally, the accuracy of typical edge detection algorithms are affected by the CCD and the lens modulation transfer function (MTF), and therefore the transition cannot be detected beyond pixel resolution [13]. However, methods have been proposed which use subpixel feature localization algorithms in order to reach super-resolution performance with low-cost CCD cameras. Examples of such techniques, together with methods for estimating edge localization errors, are presented in [66, 43, 50].

1.2.3. Popular calibration patterns. Figure 4.6 presents some common calibration patterns encountered in the literature.

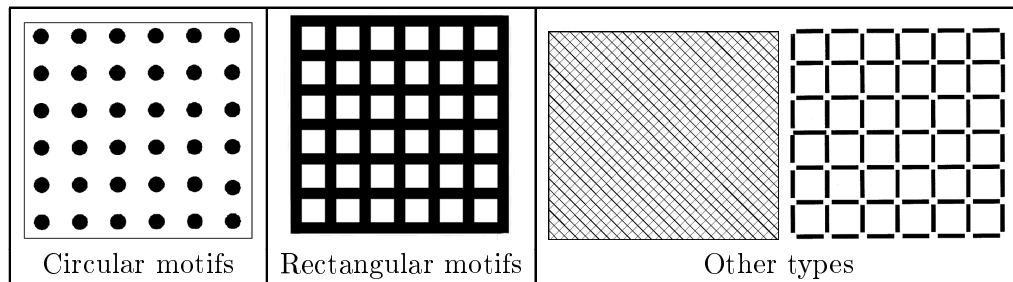


FIGURE 4.6. Example of camera calibration patterns.

Circular motifs are generally used with blob detection algorithms, but can also be used with methods combining edge detection and ellipse fitting. Rectangular motifs and other calibration markings are almost exclusively used with edge detection and/or morphological approaches. During the design phase of a calibration array, efforts have to be made to

ensure that optimal results will be achieved in terms of feature location accuracy and the resolution of the calibration points.

1.3. 2D versus 3D calibration objects. At some point during a typical calibration, accurate 3D information of the visual scene will be necessary to fix the parameters recreating the perspective projection model. This information is readily available for 3D objects through a good survey of the different markings present on the object. For 2D objects, translation and rotation stages are used to recreate the 3D information by repositioning the calibration plane after each data take. It is debatable as to which is preferable, both having advantages and disadvantages.

3D calibration objects have a shorter calibration time since only a single scene of the calibration object is required. However, accurate 3D objects are difficult to produce and maintain. Also, calibrating cameras with a large field-of-view and obtaining uniform lighting conditions is far more difficult with a 3D calibration object than a planar array.

Planar calibration tiles with appropriate surface markings can be both, more easily fabricated, and more accurately measured, compared to a 3D object. Since good quality translation and rotation stages are available at affordable prices, accurate positioning of the calibration plane in 3D space is easily attainable. Accuracy considerations for selecting rotation and translation stages should comply with the tolerance model presented in Section 1.1.

Moving a calibration tile is obviously more time consuming than taking a single image sequence of a static 3D object. However, the calibration tile approach can support all types of camera calibration methods studied for this thesis. The methods studied either require the *calibration board and the image plane to be parallel*, or simply to use the *calibration board to build a 3D object-space*. As explained in the following subsections, a calibration board mounted perpendicular to a translation stage (called the X-motion rail) serves both purposes very well.

1.3.1. Calibration board and image plane are parallel. For the generic calibration methods studied and presented later in Chapter 5, the first step of the calibration procedure is to position the calibration board perpendicular to the camera optical axis. If the optical center is determined *a priori* and the calibration board is set perpendicular to the X-motion rail, this can easily be achieved. Hence, the configuration guarantees that aligning the

center feature of the calibration board at different locations on the X-motion rail with the camera optical axis, also places the calibration board parallel to the image plane (which is perpendicular to the optical axis). For this method to work, tight tolerances on the physical alignment of the board on the X-motion rail have to be specified. Otherwise, any alignment error of the board will translate into a “pitch” or “yaw” error in the distortion correction data.

Since the alignment of the calibration board with the X-motion rail requires a great deal of mechanical precision and because the alignment is prone to drift with time, other alignment methods, using direct visual cues, were proposed. One such method, called *Mirror alignment*, is described in [47]. This method uses the camera to be calibrated and a mirror placed at the center of a pivoting calibration board (yaw and pitch). The calibration board is assumed to be perpendicular to the optical axis when the reflected image of the lens’ front ring is centered on the optical center, assuming that the optical axis passes through the center of the physical lens. Figure 4.7 depicts a typical camera scene where the calibration board is considered perpendicular to the optical axis. The mirror, attached to the calibration board, is reflecting the image from the camera body and the lens’ front ring. The alignment is deemed acceptable when the optical axis (shown with the crosshair overlay) is centered on the circular image of the lens.

As we will show, this alignment method is inaccurate for low-cost CCD cameras due to imperfections in the construction of camera housings and deficient CCD alignment methods.

Figure 4.8 illustrates the first case where some error is inserted by choosing a camera optical axis that differs from the lens optical axis. This case represents the decentering of the aperture of the lens with respect to the CCD, and the impact on the alignment of the calibration board with the X_C axis. The figure contains the light path, in a pinhole camera model, originating at the edges of the lens physical perimeter (R), reflecting in a mirror on the calibration board (h), and returning back onto the image plane (d_1 and d_2). Variables X_C and X'_C represent the lens optical axis and the incorrectly defined optical axis, respectively. The distance between the two axes is expressed by the variable d_3 . As always, the effective focal length and the distance between the calibration board and the optical center are expressed by f_e and X_{CT} , respectively. In the perfect scenario, where the calibration board is perpendicular to the optical axis and the image of the lens body is

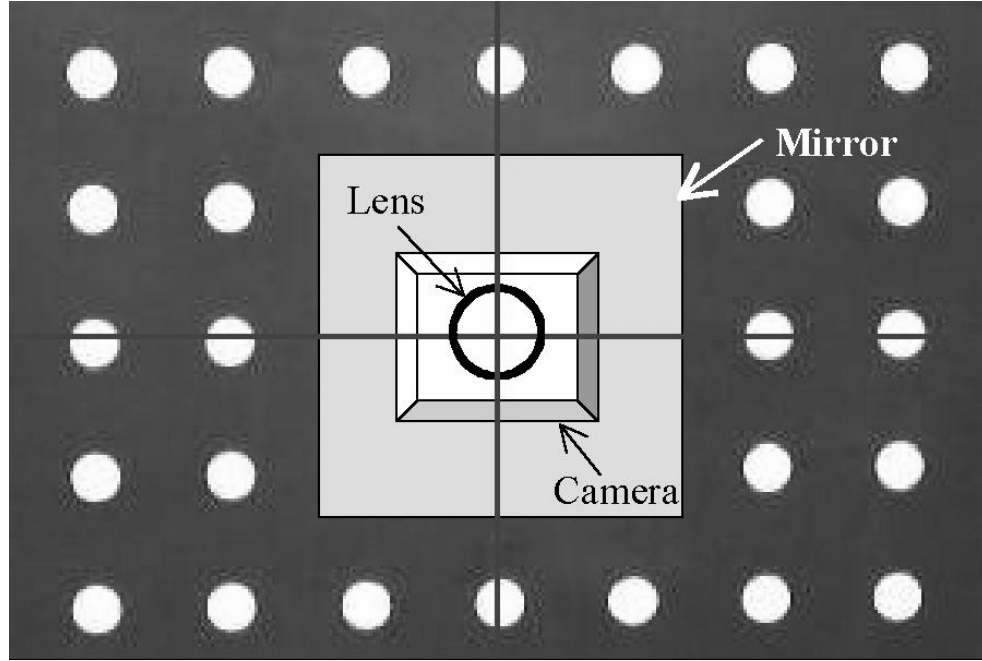


FIGURE 4.7. Simulated view from the mirror alignment technique.

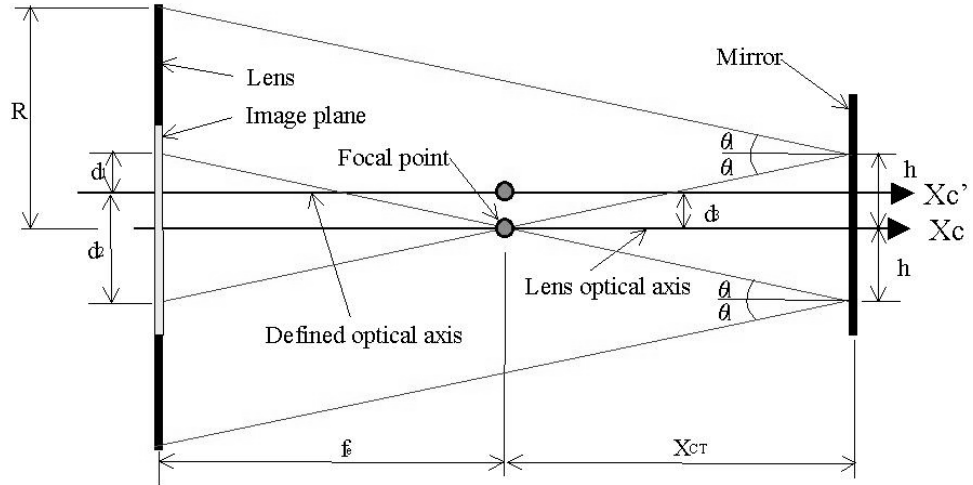


FIGURE 4.8. Calibration board alignment error (case 1).

centered on the defined optical center, the value for d_1 should be equal to d_2 . Any deviations from this rule would automatically mean that one of the alignment criteria has not been met.

In Figure 4.8, the tangent of the reflection angle (θ_1) on the mirror is given by:

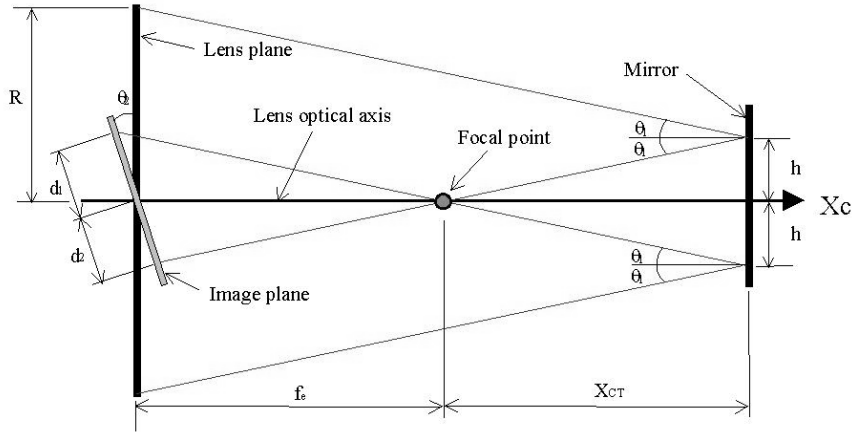


FIGURE 4.9. Calibration board alignment error (case 2).

$$(4.5) \quad \tan \theta_1 = \frac{R}{X_{CT} + f_e} = \frac{h}{X_{CT}} = \frac{d_1 + d_2}{2f_e}.$$

The values for d_1 and d_2 can be retrieved using the following substitution,

$$(4.6) \quad d_1 = d_2 - 2d_3.$$

The only condition where d_1 and d_2 are equal is when d_3 is equal to zero.

Figure 4.9 illustrates the second case where some error is inserted by a faulty CCD alignment inside the camera by an angle (θ_2). This case represents an angular misalignment of the CCD chip inside the camera housing, and the impact on the alignment of the calibration board with the X_C axis. The description of the parameters for the first case holds.

Similarly to the first case, the tangent of the reflection angle (θ_1) on the mirror is given by:

$$(4.7) \quad \tan \theta_1 = \frac{R}{X_{CT} + f_e} = \frac{h}{X_{CT}} = \frac{d_1 \cos \theta_2}{f_e + d_1 \sin \theta_2} = \frac{d_1 \cos \theta_2}{f_e - d_2 \sin \theta_2}.$$

The values for d_1 and d_2 are given by the following formulas:

$$(4.8) \quad d_1 = \frac{Rf_e}{(X_{CT} + f_e) \cos \theta_2 - R \sin \theta_2},$$

$$(4.9) \quad d_2 = \frac{Rf_e}{(X_{CT} + f_e) \cos \theta_2 + R \sin \theta_2}.$$

The only condition where d_1 and d_2 are equal is when θ_2 is equal to zero.

1.3.2. *Calibration board is used to build a 3D object-space.* Other methods studied as part of this thesis require the exact locations of calibration points in 3D space. Naturally, the perpendicularity condition of the calibration on the X-motion rail allows one to directly obtain the Y and Z target coordinates from their design locations on the calibration board. The X locations are obtained from moving the board in very precise increments along the X-motion rail, starting at the reference position T_R .

2. Camera configuration

As the first step, a complete camera calibration should include the characterization of the camera functions. During this early stage, malfunctions or defects which could affect the performance of the camera in the context of an operation are identified and corrected. A characterization normally includes a series of tests which will determine the quality of various functions inside the camera. The camera characteristics tested are closely related to the theories introduced in Chapter 2.

The following list gives an overview of some common tests used to assess the quality of a typical CCD camera.

1. *NTSC video standard* - NTSC is an acronym for National Television Standards Committee. NTSC is also commonly used to refer to composite video, RS-170A, and baseband video, which are all variations of the same thing. The NTSC standard, defined in [3], is the analog video format used throughout North America which specifies all frequencies and signal amplitudes needed to produce video on a TV monitor. All regular TV monitors or video digitizers are designed to interpret and lock on to these signals. When the video output signal from a camera does not fully comply with the standard, image feature locations and color are affected due to the inability of the digitizer to fully reconstruct the original information. The quality

of the camera output can hardly be assessed by just looking at a conventional TV monitor. More sophisticated equipment exists, such as video oscilloscopes and vectorscopes, which can verify and quantify the compliance of a signal to the standard. The most common tests to evaluate NTSC signals are described in [54].

2. *Dynamic range* - this parameter defines the range of lighting intensities at which a sensor can produce a discernable image. At high light intensities, even before smearing and blooming occur, typical CCD sensors have a tendency to bias object centroids. The phenomenon is due to an uneven CCD response to sudden luminous transitions between dark and bright regions in the scene, also called “rise-time” and “fall-time”. If uncharacterized, these centroid shifts can have detrimental effects on the accuracy of a vision system.
3. *Dynamic resolution* - there is no conventional measurement method for this parameter. However, lag effects can be measured with the freeze frame function of a video oscilloscope simulating instant transitions between dim and bright. This limitation should be considered for vision systems tracking moving objects since the lag effect shifts the target centroids in their opposite direction of travel.
4. *CCD defects* - small physical imperfections on CCD sensors are frequent in low-cost cameras. Signs of defects include dead pixels and scratches on the surface of the CCD. A popular method to identify the defects is to remove the camera lens and shine a white light directly at the CCD without saturating it. Stationary dark regions on the TV monitor generally indicate CCD defects. Significant CCD defects can interfere with the feature detection process of a vision system.
5. *Noise levels* - the most relevant sources of noise in CCD cameras are “dark current” noise, “reset” noise, and “amplifier” noise. If a CCD is operated under normal temperature conditions, the noise effects on centroids accuracy can be neglected. However, characterizing noise levels under normal operating conditions can be useful to measure stochastic imaging errors which are used by minimization methods to provide an accuracy estimation on the computed parameters. Methods for estimating CCD noise levels are provided in [29] and [8].
6. *Other camera functions* - many cameras sold these days are equipped with various electronic image control functions which extend the apparent dynamic resolution

of CCDs. Such functions include electronic shutters, automatic gain controls and gamma corrections. Their impact on centroid measurements is not always obvious and can even be detrimental. Simple tests can be conducted to assess the usefulness of these features.

Inconveniently, the applicability of a given set of camera parameters is more or less limited to the camera configuration and the lighting conditions that were selected at the time of the calibration. In an operational context, if the camera has to be used in a broad range of conditions, a wide series of calibrations has to be performed and complicated interpolation schemes between the discrete calibration points have to be considered. For guidance on this subject, see [67, 29] and their respective references.

As explained in [37], to obtain the best resolution from a lens, the iris should be set two or three stops down from the largest aperture. Non detrimental camera adjustments should then be made to obtain an optimal signal-to-noise ratio for the calibration image features. Lighting conditions are set to nominal operation levels.

For adjustable zoom and focus lenses, the repeatability of the calibration configuration is highly desirable, as each rotational lens component has to return to its calibration position to obtain the correct field-of-view and optimal distortion correction during the operation. The lack of repeatability might force the lens to be used at a hard-stop location or to be fixed mechanically.

Brand *et al.* in [10], report that the change of lens aperture has an impact on the distortion which is proportional to the lens T-number. The authors explain the phenomena by assuming that for wide lens aperture, the light rays enter from all over the lens, including the outer region which is more susceptible to radial distortion. When the iris is kept at a low aperture stop, the light rays are forced through the center of the lens, producing a minimum of distortion. This theory somewhat contradicts the material presented in Chapter 2, since geometric lens distortions are by theory independent of the lens aperture (see [60] and [42]). The aberrations reported may have been from other sources, which is typically observed with lower quality lenses. To decide if changing the iris has an impact on distortions, a simple laboratory test should be conducted where the vision system is used to track the change in centroids as the iris is stepped down. If the centroid shift is important, a solution

would be to fix the lens iris and rely on the camera's electronic shutter to maintain the average scene intensity constant.

As temperature plays such a critical role in the CCD photo-response, it is recommended that the camera run for a sufficient period of time before testing, allowing the temperature to reach a nominal operation level. Although technical studies have shown that thermal transition also affects lens optical properties, such as focus and asymmetric aberrations (see [24]), one can only trust that those conditions are minimized during an operation. Obviously, the best solution is to avoid subjecting the front of the camera to direct heat sources and hence minimizing the thermal stress.

3. Laboratory conditions

As previously mentioned, it is preferable to maintain lighting conditions for the calibration similar to conditions during the operation. In such an instance, the camera settings are optimized for a given scene contrast, better calibration data is used, and no undesired CCD artifacts are introduced. For cameras intended to be used on orbit, matching lighting conditions is a problem. On orbit lighting conditions are known for their rapid intensity change, going from full sun illumination at 10,000 foot-candles to full dark in just a few minutes. The nature of the light constantly varies as the light is coming directly from the sun or reflected from the surface of the Earth (albedo). Nevertheless, a judicious combination of a narrow-band filter and a synthetic light source can help overcome these variation problems by limiting the spectrum content of the light source.

When broadband filters are used in the camera and a sun simulator is not available, acceptable calibration results can still be obtained by using light measurement instruments and information about the camera, the scene and the light source. Some basic radiometric concepts can be applied to get the approximate amount of total light irradiance (W/cm^2) on the CCD. The formula is given by:

$$CCD_{Irradiance} = \sum_{\lambda} S(\lambda)\rho(\lambda)F(\lambda).$$

In the previous equation, $S(\lambda)$ (W/cm^2) is the light source spectral irradiance, $\rho(\lambda)$ (%) is the spectral reflectivity of the scene and $F(\lambda)$ (%) is the filter spectral transmittance value.

The concept behind the task is to adjust the alternate light source to match the expected total irradiance calculated with the intended light source. In order to maintain the color temperature of the light source, it is recommended to reduce the amount of light reaching the target by using neutral density filters or by changing the distance between the light source and the calibration board. Simple neutral density filters can be made out of glass diffusers or by layering metallic screens. We also recommend using DC powered light sources or an AC line stabilization system for conventional lamps.

Adjusting the light uniformity on the calibration board should also be performed as part of the calibration. This task can be done by measuring the image intensities from the frame grab of a video sequence or ideally by using a light measurement device located next to the camera.

The last aspect worth mentioning is to monitor the vibration level in the laboratory during the calibration. Small accelerometers can be purchased to do this task. Another option is to rest a laser pointing device on the camera and aim it on a distant wall. Monitoring the movement of the laser mark on the wall can give an indication of the level of vibration endured during the calibration or detect any camera sag.

4. Data recording

The software used for the calibrations should be flexible enough to allow the following options and information to be recorded for later analyses of calibration performance:

1. Automatically scanning different types of arrays.
2. Allow a variable number of frames to be recorded.
3. Record all measured centroids for all the target images.
4. Record all measured areas, background means, foreground means for all the target images.
5. Record any additional information such as distance from the camera, focal length, SHC vector, hardware configuration S/N, etc.
6. Stop scanning if errors are encountered.

Note that the centroid accuracy of the calibration points is normally improved if a relatively large number of consecutive frames are taken and averaged. Averaging many frames better approximates the true centroids by reducing stochastic noise and also takes advantage of the camera jitter effect that randomly shifts sampling points on the analog video signal.

5. Calibration setup used for this thesis

The last section of this chapter presents the test equipment used for our analyses.

5.1. Camera description. Three different types of cameras were chosen for this study. The first two are flight cameras, according to the NASA terminology, as they are used during Space Shuttle flights. The two flight cameras tested are the closed circuit solid state Intensified Television Camera (ITVC) and the closed circuit solid state Color Television Camera (CTVC). A Cohu monochrome camera, model 1910, with two different fixed focal length lenses was also used in this study.

The CTVCs and ITVCs are very versatile cameras due to their adjustable zoom lenses, their automatic signal gain adjustments and also because they can interface with pan/tilt units on board the Space Shuttle. Because of these advantages, they have been used the most throughout the history of SVS and will still be needed for years to come.

Some specific ISS assembly operation configurations may require NASA to install commercial off-the-shelf cameras with fixed focal length lenses at strategic locations. Most of these cameras, because of their photo-optical properties, will exhibit different characteristics than the CTVCs and ITVCs. As a result, a Cohu monochrome camera was included.

Table 4.1 contains the fundamental information about the cameras used, although Appendix A covers more detailed specifications obtained on the cameras and lenses.

	f_e	Calculated HFOV ¹	CCD size
CTVC	13.5 mm	41.15°	2/3"
ITVC	12.4 mm	40.34°	2/3"
Cohu (6 mm)	6.0 mm	55.75°	1/2"
Cohu (8 mm)	8.0 mm	43.27°	1/2"

TABLE 4.1. Fundamental camera characteristics.

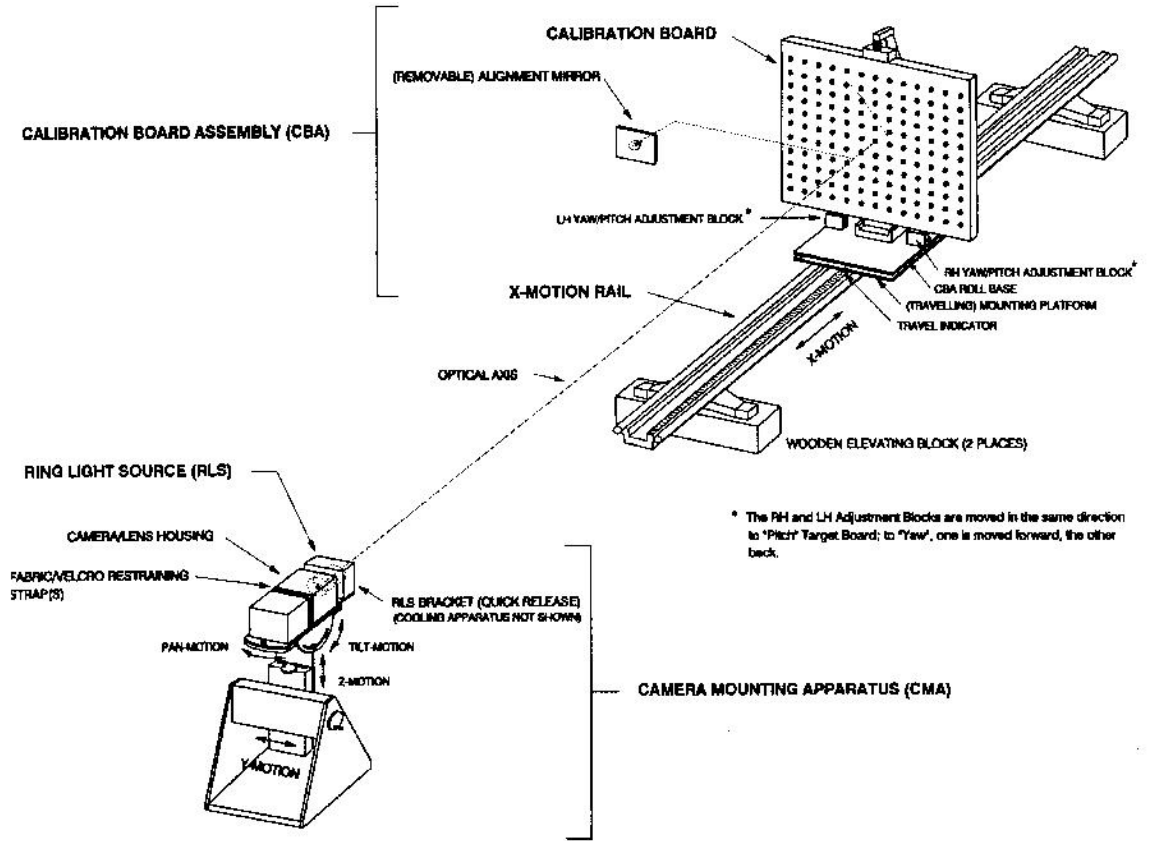


FIGURE 4.10. Camera calibration rig used for this study.

5.2. Calibration rig. Figure 4.10 contains a representation of the calibration rig used for this study. As discussed before, the Camera Mounting Apparatus (CMA) is used to securely support and control the position and orientation of the camera. It has the capability of panning, tilting and translating along the Y and Z directions. The CMA is equipped with the *Ring Light Source* (RLS), providing an adjustable and even DC illumination on the calibration object. The Calibration Object (CO) is renamed *Calibration Board Assembly* (CBA) due to our decision to use a planar calibration array mounted on a six-foot translation rail. The calibration board was precisely fixed perpendicular to the X -motion stage using precision instruments.

Figure 4.11 contains an illustration of the calibration board used for this study. It consists of a flat, black painted surface (28.5 inches by 21.5 inches) with 165 white circular 0.62-inch dots, 1.75 inches apart and arranged in an 11 by 15 array.

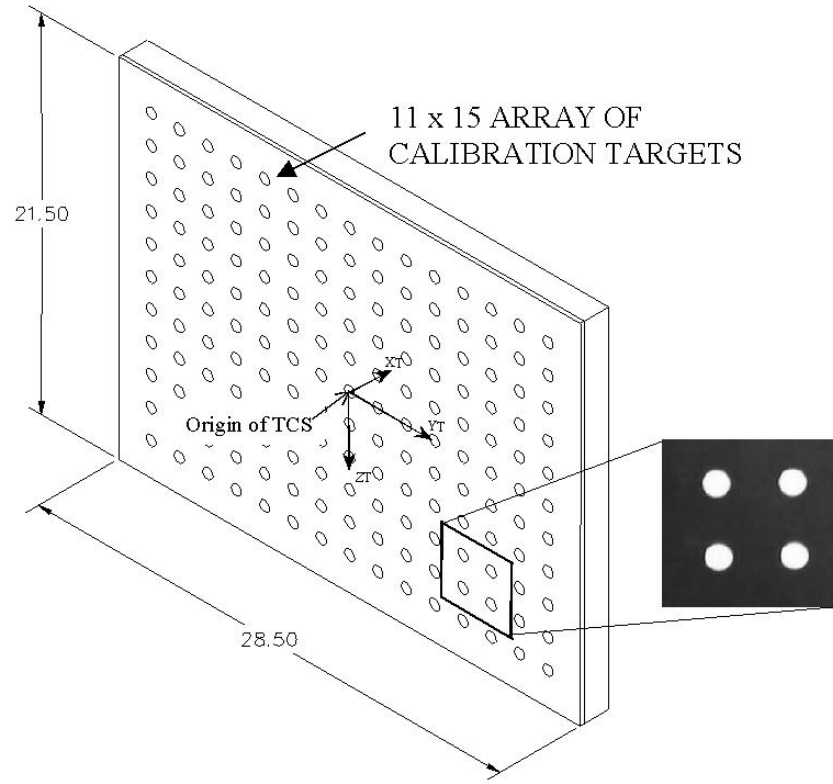


FIGURE 4.11. Calibration board used for this study.

The TCS is located on the calibration board at the geometric center of the target array, i.e., having its origin coincident with the center target. The X_T axis is perpendicular to the calibration board surface and is pointing away from the camera. The Y_T axis is aligned with the center row of targets and points right. The Z_T axis is aligned with the center column of targets and points downward.

5.3. Data collection and processing. The data used for this study was collected in the Vision System Certification Laboratory (VSCL) at Neptec Design Group Inc. The VSCL is an immense test facility primarily used for demonstrating the CSVS performance with replicas of flight hardware in simulated flight conditions. A small area of the facility is strictly dedicated to performing and studying camera calibration.

CSVS hardware version 3 and CSVS software version 5c were used during the testing. A special CSVS database file [58] specifically designed for camera calibration was used for gathering the data during the tests. A utility called *Camera Lens Calibration Utility* or

CLCUTIL exists to perform the data gathering described in Section 4 for all target elements on the calibration board. It produces an ASCII file which can easily be used by other Microsoft Windows utilities. A different data logging approach had to be used for recording the target locations for the volumetric array. The volumetric array is further explained in Chapter 6. Due to flight requirements, CSVS is constantly recording numerous parameters, including target centroids, even though the system may or may not be producing a photo-solution. The data logging utility is called *Integrated Data Logging* or IDL. By alternatively selecting all the targets of the volumetric array, a data file with all the centroids was created.

All data files from CLCUTIL or IDL were subsequently transferred from CSVS to a PC for further processing via the RS-422 interface. A copy of the complete set of test procedures, test records and data is kept in the Neptec Quality Control Database (Jeeves).

Using the same data files from CSVS, offline calibration programs were written using Matlab and Microsoft Excel. The data files from CSVS were transformed to comply with the required input format of the different applications. Since the same data was used by all calibration methods, an unbiased relative performance analysis was obtained.

6. Conclusion

General considerations on designing camera calibration rigs and on controlling the calibration conditions were presented in this chapter. The following chapter contains a description of all the methods implemented and tested as part of this evaluation exercise.

CHAPTER 5

Camera calibration methods tested

This chapter unveils the different calibration methods implemented and tested for this study. All methods can be classified as either a *generic* or a *parametric* calibration technique. The term *generic* was chosen to reflect a brute force approach where the camera parameters are derived solely based on a precise alignment of the calibration board and on a one-to-one correspondence between the distorted centroids and the ideal projection of the calibration targets. The current CSVS calibration method and a series of new ideas attempting to improve the accuracy of the original method are part of this category. At the opposite end of the spectrum, the *parametric* calibration approach is based upon the physical principles of the projection optics, camera, and data acquisition process. Standard minimization techniques are used to recover the intrinsic and extrinsic camera parameters from overdetermined sets of nonlinear equations. A description, together with implementation details of five different generic and nine parametric calibration approaches are presented in this chapter.

1. Generic calibration methods

The generic methods get their name from the approach used to rectify the distorted centroids, i.e. no distortion model is assumed therefore only a correction map, evenly spaced on the entire image plane, is produced. The correction for measured centroids falling between the corrected reference points is handled by applying a double interpolation technique using four corrected neighbors around the point to be corrected. The generic calibration methods rely on linear approximations and curve fittings to compute the distortion map.

Simple sets of linear equations based on the camera geometry are used to derive the rest of the extrinsic and intrinsic camera parameters.

Several generic calibration methods are present in the literature, but we had to select the ones that included distortion correction methods and a closed-form process to derive all the camera intrinsic parameters. Methods fulfilling these conditions were: [26], [27], [12], [15], [44], [46], and the current CSVS camera calibration method presented in [47]. However, similarities between the techniques to correct for the lens distortions were such that comparable results were expected from all of them. For all the calibration methods mentioned above, comparable criteria are used to re-establish the calibration board target geometry, using the distorted raw centroids. We verified this hypothesis by comparing the current CSVS method, which is similar to [44], against [46] and [12]. The final parameters were different, but the global calibration *self-consistency* yielded almost identical results. Calibration self-consistency for generic methods means that a perfect coupling exists between the distortion correction map and the camera intrinsic parameters. When one data set changes, the other will automatically adjust to produce the same global accuracy. This phenomenon is further explained in Section 1.3.

Our approach was then to implement an offline version of the current CSVS method, since we had to include it in our testing, and look for ways to improve it. The different approaches are described in the following subsections.

1.1. Original method. Presented in this section are the mathematical details from the current camera calibration method, as developed by the National Research Council of Canada (NRCC), and currently used to calibrate all cameras used with CSVS.

The calibration process treats the camera like a black box and forces the measured target centroids to comply with the photogrammetry colinear equations. It also assumes that the information gathered during the calibration will be self-consistent when the camera is placed in a different environment.

As seen in the Figure 4.10, the CMA provides the necessary adjustments so that the “center dot” target board array remains “centered” on the camera X_C axis during the displacement of the calibration board along the X-motion rail.

Since the calibration array has already been made parallel to the X-motion rail, once the camera optical axis is in alignment with the center dot, the calibration board is automatically perpendicular to the X_C axis.

With the calibration board perpendicular to the X_C axis, the “corrected” centroids corresponding rows and columns of dots which lie along straight lines in object-space will also lie along straight lines in the corrected image-space. Similarly, since the dots in the object-space have uniform spacing, the corrected centroids will have uniform spacing in the corrected image-space. These two conditions (consistent with the colinear equations) are implicit in the equations for deriving the true dot image positions.

To summarize the concept described above, and following the terminology in [47], the main steps performed as part of the original camera calibration method are:

1. By successive displacement of the calibration board on the X-motion rail, the CMA is positioned and oriented until the pre-defined camera optical axis is well aligned with the “center dot” of the calibration board at any location on the X-motion rail. For CSVS, the optical center is defined in image-space by the nominal center of the NTSC picture area. The value normally used is pixel 373 and line 240.
2. When the calibration board is brought to the near end of the X-motion rail (i.e. the calibration board fills the camera field of view), the CSVS is used to record the centroid (y_D, z_D) of the 165 (11 by 15) uniformly spaced target elements of the calibration board. This location becomes the reference calibration location designated by T_R .

The data collected during step 2 is now used to derive the correction map and preliminary gain ratios. Later, this data (reference position of calibration array) is combined with a second set of data taken at another position of the calibration array to compute the remaining calibration parameters (SHC , refined gain ratios, and final correction map). The camera parameters are obtained starting with the next step.

3. As mentioned before, the nominal value of the optical center is 373 pixels and 240 lines. However, the limited accuracy of the alignment process means that the measured position of the center dot is not exactly equal to the nominal value of the

optical center. The position of the “true” array is derived from the measured centroid of the center dot after alignment. Row i and column j of the center dot are given by:

$$\begin{aligned} i &= \frac{N_{row} + 1}{2}, \\ j &= \frac{N_{col} + 1}{2}, \end{aligned}$$

where N_{row} and N_{col} are respectively the number of rows and columns on the calibration board.

The center dot centroid, being used as the new optical center, is given by:

$$\begin{aligned} C_y &= y_D[i, j], \\ C_z &= z_D[i, j]. \end{aligned}$$

It is assumed that the calibration array has an odd number of rows and an odd number of columns.

4. Using the following formulae, it is possible to measure the average dot spacing of the true array by averaging the distance between selected points on the calibration array.

$$\begin{aligned} row_{\Delta} &= \sum_{i=2}^{N_{row}-1} ((y_D[i, N_{col} - 1] - y_D[i, 2]) / (N_{row} - 2)(N_{col} - 3)), \\ col_{\Delta} &= \sum_{j=2}^{N_{col}-1} ((z_D[N_{row} - 1, j] - z_D[2, j]) / (N_{row} - 3)(N_{col} - 2)). \end{aligned}$$

The previous row and column selection is arbitrary. A subset of the points is used to obtain a better value of row_{Δ} , col_{Δ} over the central region (70%) of the image where non-linear contributions are typically small. Accordingly, N_{row} and N_{col} should be greater than or equal to five to satisfy the basic requirements for calibration.

5. The rotation of the calibration board with respect to the image plane is also used for deriving the true dot position. The information is gathered according to the following steps:

- (a) CSVS measures the slope of the nominally vertical columns (m_y) and nominally horizontal rows (m_z) using:

$$\begin{aligned}
 m_y &= \frac{6}{(N_{row} - 1)N_{row}(N_{row} + 1)col_{\delta}} \cdot \sum_{i=1}^{N_{row}} (2i - N_{row} - 1) \\
 &\quad \cdot \frac{1}{N_{col} - 2} \sum_{j=2}^{N_{col}-1} y_D[i, j], \\
 m_z &= \frac{6}{(N_{col} - 1)N_{col}(N_{col} + 1)row_{\delta}} \cdot \sum_{j=1}^{N_{col}} (2j - N_{col} - 1) \\
 &\quad \cdot \frac{1}{N_{row} - 2} \sum_{i=2}^{N_{row}-1} z_D[i, j].
 \end{aligned}$$

- (b) The new slope coefficients (n_y, n_z) are calculated to relate the target spacing on the calibration board to the centroid slopes measured using the next equations:

$$\begin{aligned}
 K &= \left(\frac{Brow_{\Delta}}{Bcol_{\Delta}} \right) \cdot \left(\frac{col_{\Delta}}{row_{\Delta}} \right), \\
 a_y &= -K \cdot m_y, \\
 a_z &= -\frac{1}{K} \cdot m_z,
 \end{aligned}$$

where $Brow_{\Delta}$ and $Bcol_{\Delta}$ are the distance between target rows and columns respectively on the calibration board.

The final estimate of the rotation (a) coefficient is the average of the two slope angles a_y and a_z and is given by:

$$a = \frac{a_y + a_z}{2}.$$

The final slope coefficients are given by:

$$\begin{aligned} n_y &= \frac{a}{K}, \\ n_z &= aK. \end{aligned}$$

6. The true dot image positions ($T_y[i, j]$, $T_z[i, j]$) can be calculated using:

$$\begin{aligned} T_y[i, j] &= C_y + \left(j - \left(\frac{N_{col} + 1}{2}\right)\right) \cdot col_{\Delta} + \left(i - \left(\frac{N_{row} + 1}{2}\right)\right) \cdot row_{\Delta} \cdot n_y, \\ T_z[i, j] &= C_z + \left(i - \left(\frac{N_{row} + 1}{2}\right)\right) \cdot row_{\Delta} + \left(j - \left(\frac{N_{col} + 1}{2}\right)\right) \cdot col_{\Delta} \cdot n_z. \end{aligned}$$

7. The final correction matrices ($Corr_y$ and $Corr_z$) components are obtained using:

$$\begin{aligned} Corr_y[i, j] &= T_y[i, j] - y_D[i, j], \\ Corr_z[i, j] &= T_z[i, j] - z_D[i, j]. \end{aligned}$$

The details in regard to calculating the final correction factors D_y and D_z are presented in Section 1.1.1, but note that the correction matrices are used in conjunction with specific locations in the image plane. For convenience, the correction factors are attributed to the true dot image positions ($T_y[i, j]$, $T_z[i, j]$) rather than the measured dot image position ($y_D[i, j]$, $z_D[i, j]$). Following this approach, the entire location matrix can be derived from four parameters: $T_y[1, 1]$, $T_z[1, 1]$, row_{Δ} , and col_{Δ} .

8. Using the distorted centroids from the calibration board targets obtained at different locations on the X-motion rail, and the estimated effective focal length (f_e), it is now possible to obtain the gain ratios (G_y , G_z) defining the size of the CCD photosensitive elements.

$$\begin{aligned} G_y &= \frac{Bcol_{\Delta}}{Dist} \cdot \frac{f_e}{col_{\Delta}}, \\ G_z &= \frac{Brow_{\Delta}}{Dist} \cdot \frac{f_e}{row_{\Delta}}, \end{aligned}$$

where $Dist$ is the final calibrated distance in inches between the calibration board at location T_R and the camera optical center.

Note that the value for $Dist$ is obtained by comparing the two photosolutions and the physical displacement obtained by moving the calibration board on the X-motion rail by about 45 inches.

1.1.1. *Run-time use of the correction map.* As mentioned in the previous subsection, the original calibration method returns the final correction components ($Corr_y$ and $Corr_z$) in a N_{row} by N_{col} matrix form. The location matrix concept is introduced to express the true location of the various correction map elements in the image plane. Four calibrated terms ($T_y[1, 1]$, $T_z[1, 1]$, row_Δ , and col_Δ), as well as the index of the row and column of a specific element, are necessary to determine the reconstructed centroids. The true locations are given by:

$$\begin{aligned} T_y[i, j] &= T_y[1, 1] + (j - 1) \cdot col_\Delta, \\ T_z[i, j] &= T_z[1, 1] + (i - 1) \cdot row_\Delta. \end{aligned}$$

The corrections in the correction map are associated with the true dot locations.

Suppose (y_F, z_F) is a measured point to be corrected. The first step is to use the location matrix to find the pair of rows $i, i + 1$ and the pair of columns $j, j + 1$ which define the rectangle within the correction matrix which is closest to the point (y_F, z_F) . The second step is to determine the interpolation coefficients a and b , used for interpolation in y and z respectively. The correction to apply to (y_F, z_F) is defined by interpolation or extrapolation using the four corrections at the corners of the rectangle.

Strictly speaking, the measured point (y_F, z_F) should be corrected by relating it to the original measured points used to define the correction matrix, not to the true positions. The shortcut used is an approximation that has been considered acceptable since the measured and the true points are usually not too far apart.

To determine the column index j and horizontal interpolation coefficient a , an intermediate quantity FJ , the fractional distance, is calculated. The fractional distance is a real

number. The integral part of FJ is $j - 1$ and the fractional part is the interpolation coefficient. The fractional distance may be interpreted as the distance in pixels of the measured point from the left-hand edge of the correction matrix ($T_y[1, 1]$). The sequence of operations is the following:

$$FJ = \frac{(y_F - T_y[1, 1])}{col_{\Delta}}.$$

If FJ is less than or equal to 1, then set $j = 1$.

If FJ is greater than or equal to $N_{col} - 2$, then set $j = N_{col} - 1$.

Else, set $j = INT(FJ) + 1$ (where INT is the integer part of “function”).

The interpolation coefficient a is finally defined by: $a = FJ - (j - 1)$.

Similarly, the row index i and the vertical interpolation coefficient b are defined in terms of the vertical fractional distance FI , as follows:

$$FI = \frac{(z_F - T_z[1, 1])}{row_{\Delta}}.$$

If FI is less than or equal to 1, then set $i = 1$.

If FI is greater than or equal to $N_{row} - 2$, then set $i = N_{row} - 1$.

Else, set $i = INT(FI) + 1$.

The interpolation coefficient b is finally defined by: $b = FI - (i - 1)$.

The final correction terms (D_y and D_z) to be applied are obtained by interpolation (extrapolation) of the $Corr_y$ and $Corr_z$ matrices between the points (i, j) , $(i+1, j)$, $(i, j+1)$, and $(i+1, j+1)$. The formula for the corrections to be applied is also valid when (y_F, z_F) lies outside the defined calibration map.

The value for D_y is given by:

$$\begin{aligned} K_{y1} &= Corr_y[i, j] + a(Corr_y[i, j+1] - Corr_y[i, j]), \\ K_{y2} &= Corr_y[i+1, j] + a(Corr_y[i+1, j+1] - Corr_y[i+1, j]), \\ D_y &= K_{y1} + b(K_{y2} - K_{y1}). \end{aligned}$$

The value for D_z is given by:

$$\begin{aligned}
K_{z1} &= Corr_z[i, j] + a(Corr_z[i, j + 1] - Corr_z[i, j]), \\
K_{z2} &= Corr_z[i + 1, j] + a(Corr_z[i + 1, j + 1] - Corr_z[i + 1, j]), \\
D_z &= K_{z1} + b(K_{z2} - K_{z1}).
\end{aligned}$$

The following equations are used to convert the distorted measurements in the FCS (y_F, z_F) to corrected centroids in the ICS (y_U, z_U) :

$$\begin{aligned}
y_U &= ((y_F + D_y) - C_y)G_y, \\
z_U &= ((z_F + D_z) - C_z)G_z.
\end{aligned}$$

1.2. Lookup compensation method. As explained in the previous section, a shortcut has been taken during the calculation of the indices of the distortion correction map. Hence, a correction factor measured from a distorted image point should not readily be applied to a nearby reference point used to derive the correction. The effect is benign near the center of the field-of-view where lower lens distortion exists, but may be significant near the edges of the field-of-view. *Lookup compensation* is a generic term describing techniques which determine the true corrections that should be applied to the calibration reference grid.

The method chosen for this task was first to fit a surface to the nonuniformly-spaced distorted points and then interpolate (or extrapolate) this surface at the points specified by the reference correction grid.

Various interpolation approaches were considered and tested, but the triangle-based linear interpolation method from [61] was found to be the most robust.

The main steps and equations involved in the Lookup compensation method are now given.

1. The problem is separated into two interpolation tasks. One for measuring the compensated correction terms CD_y in the y image axis and another for the compensated correction terms CD_z in the z image axis.

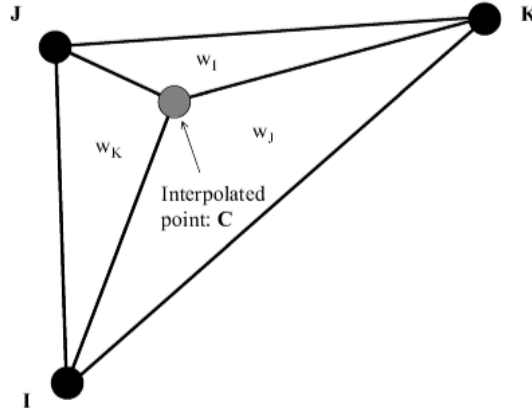


FIGURE 5.1. Barycentric local coordinates.

The measured calibration points and associated correction terms are grouped into two sets of vectors (y_F, z_F, D_y) and (y_F, z_F, D_z) , where (y_F, z_F) are the distorted centroids in the FCS and $(D_y$ and $D_z)$ are the final correction terms returned by the generic calibration method described in Section 1.1.

The true grid location terms (T_y, T_z) are the interpolation reference points used to derive CD_y and CD_z .

2. For each point i on the uniform grid (T_y, T_z) , the three closest points amongst (y_F, z_F) are found. The distance between a reference point (T_{y_i}, T_{z_i}) and a measurement point (y_{F_j}, z_{F_j}) is given by:

$$Dist_{i,j} = \sqrt{(T_{y_i} - y_{F_j})^2 + (T_{z_i} - z_{F_j})^2}.$$

3. The next step is to compute the barycentric coordinates using the formula given in [61, page 78]. Barycentric coordinates relate a given location C to three data points I, J and K, as in Figure 5.1. The areas of the subtriangles, CIJ, CJK, and CKI, each calculated as fractions of the area of the triangle IJK, are the barycentric coordinates of the interpolation point C for each opposite vertex, respectively.

The following equation is used to measure the barycentric coordinates from three measured points (y_{F_I}, z_{F_I}) , (y_{F_J}, z_{F_J}) and (y_{F_K}, z_{F_K}) for the interpolation point (T_y, T_z) :

$$\begin{aligned}
w_I &= ((y_{FJ} - T_y) \cdot (z_{FK} - T_z) - (y_{FK} - T_y) \cdot (z_{FJ} - T_z)) / Det_{IJK}, \\
w_J &= ((y_{FK} - T_y) \cdot (z_{FI} - T_z) - (y_{FI} - T_y) \cdot (z_{FK} - T_z)) / Det_{IJK}, \\
w_K &= ((y_{FI} - T_y) \cdot (z_{FJ} - T_z) - (y_{FJ} - T_y) \cdot (z_{FI} - T_z)) / Det_{IJK},
\end{aligned}$$

where Det_{IJK} is given by:

$$Det_{IJK} = (y_{FJ} - y_{FI}) \cdot (z_{FK} - z_{FI}) - (y_{FK} - y_{FI}) \cdot (z_{FJ} - z_{FI}).$$

4. The final compensated correction term CD_y and CD_z are calculated using:

$$\begin{aligned}
CD_y &= D_{yI} \cdot w_I + D_{yJ} \cdot w_J + D_{yK} \cdot w_K, \\
CD_z &= D_{zI} \cdot w_I + D_{zJ} \cdot w_J + D_{zK} \cdot w_K.
\end{aligned}$$

1.3. Averaging of multiple compensated maps. Data averaging between consistent calibration sets reduces the influence of stochastic measurement noise and further improves the level of certainty for points being interpolated in the reference grid. Ultimately, the distortion correction information from non-colocated maps can be used to increase the resolution and accuracy of the final correction map.

Special precautions have to be taken when averaging data obtained with generic methods since two calibration sets, obtained from two board positions, will very likely produce incompatible results. This is due to the disparity between the gain ratios measured using distorted target centroids taken at various locations on the image plane. The generic methods are based on the self-consistency of all the perspective projection terms which do not necessarily correspond with the “real” fixed camera parameters, which in turn could be used as a reference.

To demonstrate this point, we calibrated the Cohu camera with the 8 mm lens, selecting different columns and rows in step 3 of the original calibration method. Naturally, the values for the row_Δ and col_Δ parameters were changing as a function of the local distortions in the camera field-of-view. The parameters row_Δ and col_Δ are measured in the early stage of the

calibration and form the basic information which is used to derive the rest of the pinhole camera parameters, including the final distortion map. Figure 5.2 presents the variation of the Cohu camera gain ratios (G_y , G_z) with different combinations of lines and columns. Since the gain ratios for the generic methods are calculated using a subset of distorted centroids, a true representation of the actual CCD size can never be obtained. Hence, the gain ratios vary to accommodate the correction map and therefore should always be used together. Not surprisingly, the trend of change in gain ratios generally has a quadratic shape which agrees with the radial distortion model considered in the parametric calibration methods.

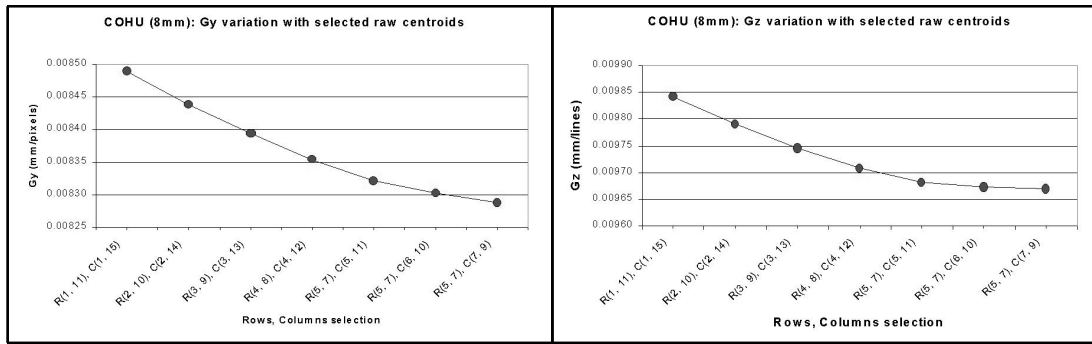


FIGURE 5.2. Variation of the gain ratios from the selection of different distorted target elements for the Cohu camera with an 8 mm lens.

A *map averaging* method which deals with the gain ratio variation problem, is now presented. The approach simply fixed certain camera parameters between different calibration sets to ensure consistency, therefore allowing the calibration maps to be combined. The main steps and equations involved in the map averaging method are as follows:

1. Perform the original calibration method following each step described in Section 1.1. Values for G_y , G_z and $Dist$ are saved for later steps.
2. For each additional data set, the relative distance of the calibration board to the first calibration location T_R has to be measured. The new variable is labeled $Dist_\Delta$. New values for the distance between rows and columns are then calculated using:

$$\begin{aligned}
Nrow_{\Delta} &= \frac{f_e \cdot Brow_{\Delta}}{(Dist_{\Delta} + Dist) \cdot G_z}, \\
Ncol_{\Delta} &= \frac{f_e \cdot Bcol_{\Delta}}{(Dist_{\Delta} + Dist) \cdot G_y}.
\end{aligned}$$

The new calibration maps are derived following steps 5 through 7 presented in Section 1.1, using $Nrow_{\Delta}$ and $Ncol_{\Delta}$ as measured above.

3. The lookup compensation method presented Section 1.2 is applied to all measured data sets, including the first calibration, to obtain the correction factors for a common evenly-spaced reference grid.
4. The last step is to derive the final correction maps by averaging, for each point of the reference grid, the correction terms obtained from all data sets.

1.4. Polynomial smoothing method. The *Polynomial smoothing* method attempts to improve the correction map obtained with the original calibration method using a polynomial data fitting approach.

The intent is to mitigate the stochastic measurement errors at the reference points by readjusting the correction terms using a smoothness transition constraint between correction terms in the vicinity of each other.

This method generates a smoothed correction map from an unsmoothed one. Correction map smoothing affects only the correction map, and is done independently of the remaining camera parameters.

Smoothing is done by a 1-dimensional polynomial fit for the corrections in the correction map. Curve fitting is done in each column for y corrections and in each row for z corrections. For this thesis, a degree four was used for all polynomials.

Smoothed y corrections are given by:

$$SD_y = a_0 + a_1(i - 6) + a_2(i - 6)^2 + \dots + a_n(i - 6)^n,$$

where a_0, a_1, \dots, a_n are the coefficients chosen to give the least-squares residuals. Smoothing is done independently for each column. The variable i represents the row number and $i = 6$ corresponds to the middle row.

Smoothed z corrections are given by:

$$SD_z = a_0 + a_1(j - 8) + a_2(j - 8)^2 + \dots + a_n(j - 8)^n,$$

where a_0, a_1, \dots, a_n are the coefficients chosen to give the least-squares residuals. Smoothing is done independently for each row. The variable j represents the column number and $j = 8$ corresponds to the middle column.

1.5. Polynomial smoothing with lookup compensation method. The *Polynomial smoothing with lookup compensation* is an extension of the Polynomial smoothing method presented in Section 1.4.

Lookup compensation is obtained by using the polynomial fitting curves to determine the appropriate corrections on the reference grid. The point at which corrections are looked up normally lies between two fitting curves. The two fitting curves are used to estimate two corrections. Linear interpolation between the two corrections is done to find the final correction. For y corrections, vertical interpolation is done with the fitting curves, and horizontal interpolation is linear. For z corrections, horizontal interpolation is done with the fitting curves, and vertical interpolation is linear.

2. Parametric calibration methods

Given a set of observations (distorted centroids), the parametric methods attempt to condense and summarize the data by fitting it to a “model” that depends on adjustable camera parameters. The models come from the underlying pinhole camera theory presented in Chapter 3, and from physical principles modeling specific aberrations. The same intrinsic and extrinsic camera parameters pursued by the generic methods are obtained with the parametric methods, but instead of a calibration map, the centroid distortions are corrected using mathematical equations with calibrated coefficients.

The various steps performed during a parametric calibration are very simple to follow. A summary of what they are is now presented:

1. The calibration board has to be precisely adjusted perpendicular to the X-motion rail. This step is crucial since the 3D object-space is reconstructed assuming that the

calibration targets are in a plane perpendicular to the axis defined by the X-motion rail direction of displacement.

2. The camera is roughly aligned with the calibration board at the initial position for all the dots to be inside the desired field-of-view.
3. The target array centroid information is collected at each step while the calibration board is gradually translated by uniform increments on the X-motion rail. The initial calibration board location becomes the reference position T_R . A total of ten different positions were used for this study.
4. The camera parameters are extracted by fitting the centroid information to a “model” underlying the observation data.

This section is mainly dedicated to explaining the different distortion models tested and the solution variants for similar models. The general approach considered for this thesis was to simultaneously solve for all camera parameters while insuring that a good initial guess was used. The following subsection includes brief details about solving the nonlinear equations modeling a distorted pinhole camera.

2.1. The solution process for the parametric methods. As for all data fitting problems, one needs to choose a figure-of-merit function that measures the agreement between the data and the model with a particular choice of parameters. The figure-of-merit function chosen here to solve the nonlinear parametric calibration problem is the Least-Squares (LS) fitting criterion. The solution first proceeds by linearizing the colinear relations and then solving, by using an iterative process, until a convergence threshold is met. The theory found in [5] and some practical details found in [55] were used in the implementation of all the parametric methods.

The colinear relations of photogrammetry from Section 2 are used to derive the mathematical expressions to minimize χ by the LS fitting criterion for a total number of N targets:

$$(5.1) \quad \chi = \arg \min_{\vec{x}} \left[\sum_{i=1}^N (\Delta_y^2(i) + \Delta_z^2(i)) \right],$$

where

$$\begin{aligned}
(5.2) \quad \Delta_y(i) &= y_U(i) (X_{CT} + X_T \alpha_{11} + Y_T \alpha_{12} + Z_T \alpha_{13}) \\
&\quad - f_e (Y_{CT} + X_T \alpha_{21} + Y_T \alpha_{22} + Z_T \alpha_{23}), \\
\Delta_z(i) &= z_U(i) (X_{CT} + X_T \alpha_{11} + Y_T \alpha_{12} + Z_T \alpha_{13}) \\
&\quad - f_e (Z_{CT} + X_T \alpha_{31} + Y_T \alpha_{32} + Z_T \alpha_{33}).
\end{aligned}$$

From the distorted centroids (y_F, z_F) in the FCS, the corrected centroids (y_U, z_U) in the ICS are obtained using the following equations:

$$\begin{aligned}
y_U &= (y_F - C_y) \cdot G_y + D_y, \\
z_U &= (z_F - C_z) \cdot G_z + D_z.
\end{aligned}$$

The distortion correction elements (D_y, D_z) are obtained using various distortion models which will be described in the following subsection. The number of unknowns in the colinear equations vary depending on the distortion models used.

A complete example of the method is given in Appendix B for a pinhole camera model afflicted with first order radial lens distortions.

Note that the method used to minimize Equation 5.1 does not provide an estimate of accuracy on the computed parameters. However, it has the advantage of being simple to implement and to converge extremely fast. Other methods, such as the *Bard-Deming algorithm* or the *Extended Kalman Filter* (presented in [17]), which take into account the measurement uncertainty to return the estimated parameter accuracy, should be used to obtain robust solutions when measurement outliers and large noise levels are present. For this study, it is understood that the centroid data was scrutinized for outliers and an average of several video frames was used to reduce the effects of noise.

2.2. Parametric distortion models. The distortion models now presented were mainly found in [53].

2.2.1. Radial distortion model. Radial distortion comes from the fact that normal lenses do not follow the pinhole lens model, and so curvature imperfections are detected. The image points are displaced along a vector having its origin located at the center of the

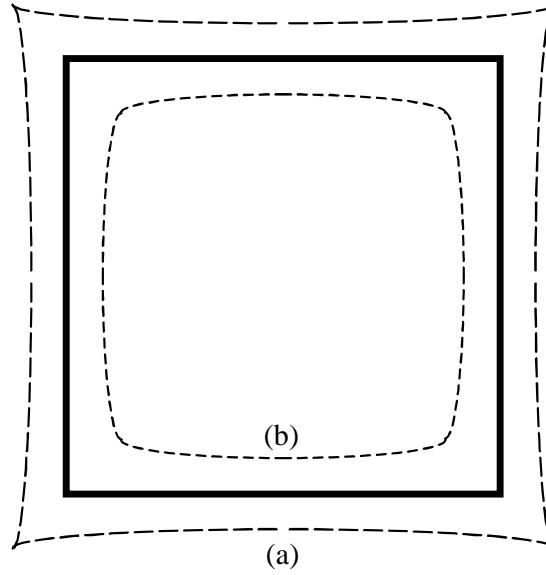


FIGURE 5.3. Radial distortion effects.

image and oriented where the ideal point would be. The norm of the vector is proportional to the distance between the point and the origin. The radial effects are shown in Figure 5.3. The lines that are continuous represent the image without distortions. Dashed lines represent the radial distortions. Positive distortions (a) will cause the off-axis points to be imaged at distances greater than nominal, creating the pincushion effect. Negative distortions (b) will produce a barrel shape with off-axis points imaged at distances smaller than nominal. The type of lens will directly influence the sign and amplitude of the radial distortions.

The radial distortion model is expressed as follows:

$$D_{rad_y} = y_D (k_1 r_D^2 + k_2 r_D^4 + \dots),$$

$$D_{rad_z} = z_D (k_1 r_D^2 + k_2 r_D^4 + \dots),$$

where: $r_D = \sqrt{y_D^2 + z_D^2}$

2.2.2. *Tangential distortion model.* Tangential distortions are the result of a poor assembly between the lens and the camera image plane. The distortions may result from

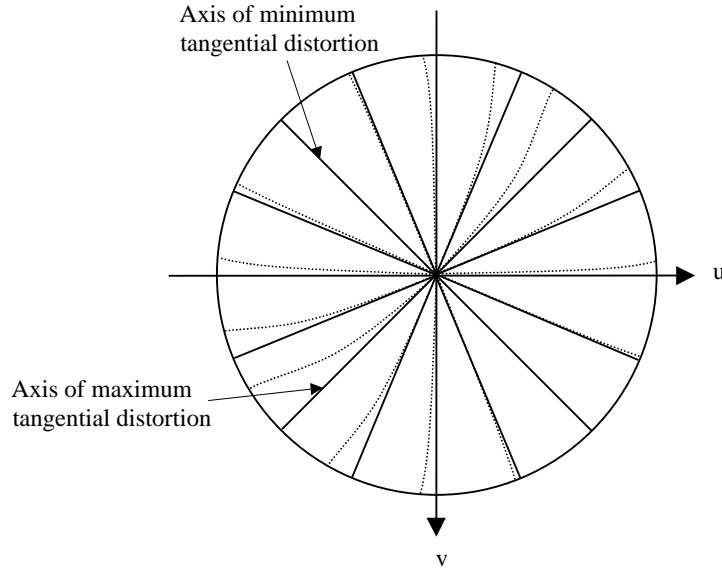


FIGURE 5.4. Tangential distortion effects.

a bad lateral alignment (decentering distortion) and/or because the lens optical axis is not parallel to the image plane's normal vector (thin prism distortion). The main effect is to displace the image points tangentially to the circle centered on the image principal point as shown in Figure 5.4. The lines that are continuous represent the image without distortions. Dotted lines represent the tangential distortions.

The tangential distortion can be modeled as follows:

$$\begin{aligned} Dtan_y &= [k_1 (r_D^2 + 2y_D) + 2k_2 y_D z_D] + k_3 r_D^2 + k_4 r_D^4 + \dots, \\ Dtan_z &= [k_2 (r_D^2 + 2z_D) + 2k_1 y_D z_D] + k_3 r_D^2 + k_4 r_D^4 + \dots \end{aligned}$$

Taken separately, the decentering and the thin prism distortions (first order only) are expressed by:

$$\begin{aligned}
Ddec_y &= k_1 (3y_D^2 + z_D) + 2k_2 y_D z_D, \\
Ddec_z &= k_2 (y_D^2 + 3z_D^2) + 2k_1 y_D z_D, \\
Dtp_y &= Dtp_z = k_3 (y_D^2 + z_D^2).
\end{aligned}$$

2.2.3. *Other distortion models.* Other distortion models can be used to duplicate more exotic distortion types (see [53]). Additional, less frequently used models are presented in this section.

A special lens having different distortion coefficients along the two axes can be modeled by:

$$\begin{aligned}
Dsrady &= k_1 y_D r_D^2, \\
Dsradz &= k_2 z_D r_D^2.
\end{aligned}$$

Second order distortion effects are typically represented by:

$$\begin{aligned}
D2ord_y &= k_1 y_D^2 + k_2 z_D^2, \\
D2ord_z &= k_3 y_D^2 + k_4 z_D^2.
\end{aligned}$$

2.3. Parametric distortion models tested. In this study, ten different parametric approaches were tested. The same distortion models are repeated between different implementations, but either the list of intrinsic camera parameters to be solved is altered or a two-step calibration is used.

The next subsections present the specific details about the different parametric models implemented.

2.3.1. *Radial - fixed focal length.* The *Radial - fixed focal length* implementation is directly inspired by the radial distortion model found in Section 2.2.1. A single order coefficient k_1 is used.

The effective focal length is assumed known so the gain ratios remain to be determined.

- Unknowns: $G_y, G_z, C_y, C_z, K_1, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R},$ and α_{23R} .

- Equations to solve: 5.1 (formed with 5.3 and 5.3), where

$$D_y(i) = y_D(i)k_1r_D^2,$$

$$D_y(i) = z_D(i)k_1r_D^2.$$

2.3.2. *Radial - fixed gain ratios.* The *Radial - fixed gain ratios* implementation is directly inspired by the radial distortion model found in Section 2.3.1 since a single order coefficient k_1 is used.

The size and resolution of the camera image plane (G_y, G_z) is assumed known so that the effective focal length and the uncertainty scale factor S_y for G_y remain to be established.

- Unknowns: $f_e, S_y, C_y, C_z, K_1, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R}$, and α_{23R} .
- Equations to solve: 5.1 (formed with 5.3 and 5.3), where

$$D_y(i) = y_D(i)k_1r_D^2,$$

$$D_y(i) = z_D(i)k_1r_D^2.$$

2.3.3. *Radial - different data sets.* The *Radial - different data sets* implementation verifies the parameters' consistency of the method found in Section 2.3.1. The approach is to compare the calibration coefficients when different input data sets are used.

2.3.4. *Radial and decentering.* The *Radial and decentering* method combines the distortion models for radial distortion and decentering found in Sections 2.2.1 and 2.2.2, respectively.

The effective focal length is assumed known so the gain ratios remain to be determined.

- Unknowns: $G_y, G_z, C_y, C_z, k_1, k_2, k_3, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R}$, and α_{23R} .
- Equations to solve: 5.1 (formed with 5.3 and 5.3), where

$$D_y(i) = k_1y_D(i)(y_D^2(i) + z_D^2(i)) + k_2(3y_D^2(i) + z_D^2(i)) + 2k_3y_D(i)z_D(i),$$

$$D_z(i) = k_1z_D(i)(y_D^2(i) + z_D^2(i)) + k_3(3z_D^2(i) + y_D^2(i)) + 2k_2y_D(i)z_D(i).$$

2.3.5. *Radial and decentering - 2 steps.* The *Radial and decentering - 2 steps* implementation is achieved by sequentially combining the distortion models found in Sections 2.2.1 and 2.2.2. During the first minimization the single order radial coefficient k_1 is set. The radial coefficient is then re-injected in a second minimization routine using the combined radial and decentering model.

The effective focal length is assumed known so the gain ratios remain to be determined.

- Unknowns: $G_y, G_z, C_y, C_z, k_1, k_2, k_3, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R},$ and α_{23R} .
- Equations to solve: 5.1 (formed with 5.3 and 5.3).

During the first minimization step, the distortion model used is:

$$\begin{aligned} D_y(i) &= y_D(i)k_1r_D^2, \\ D_z(i) &= z_D(i)k_1r_D^2. \end{aligned}$$

During the second minimization step, the radial distortion parameter (k_1) is kept constant. The new distortion model used is:

$$\begin{aligned} D_y(i) &= k_1y_D(i)(y_D^2(i) + z_D^2(i)) + k_2(3y_D^2(i) + z_D^2(i)) + 2k_3y_D(i)z_D(i), \\ D_z(i) &= k_1z_D(i)(y_D^2(i) + z_D^2(i)) + k_3(3z_D^2(i) + y_D^2(i)) + 2k_2y_D(i)z_D(i). \end{aligned}$$

2.3.6. *Radial and tangential.* The *Radial and tangential* method combines the distortion models for radial distortions and tangential distortions found in Sections 2.2.1 and 2.2.2, respectively.

The effective focal length is assumed known so the gain ratios remain to be determined.

- Unknowns: $G_y, G_z, C_y, C_z, k_1, k_2, k_3, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R},$ and α_{23R} .
- Equations to solve: 5.1 (formed with 5.3 and 5.3), where

$$\begin{aligned} D_y(i) &= [k_1(r_D^2 + 2y_D) + 2k_2y_Dz_D] + k_3r_D^2 + k_4r_D^4, \\ D_z(i) &= [k_2(r_D^2 + 2z_D) + 2k_1y_Dz_D] + k_3r_D^2 + k_4r_D^4. \end{aligned}$$

2.3.7. *Radial and thin prism.* The *Radial and thin prism* calibration implementation is similar in principle to the *Radial and decentering* coding. The thin prism distortion model presented in Section 2.2.2 now replaces the decentering model.

The effective focal length is assumed known so the gain ratios remain to be determined.

- Unknowns: $G_y, G_z, C_y, C_z, k_1, k_2, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R},$ and α_{23R} .
- Equations to solve: 5.1 (formed with 5.3 and 5.3), where

$$\begin{aligned} D_y(i) &= k_1 y_D(i) (y_D^2(i) + z_D^2(i)) + k_2 (y_D^2(i) + z_D^2(i)), \\ D_z(i) &= k_1 z_D(i) (y_D^2(i) + z_D^2(i)) + k_2 (y_D^2(i) + z_D^2(i)). \end{aligned}$$

2.3.8. *Radial and second order.* The *Radial and second order* calibration model can be obtained by combining the concepts introduced in Section 2.2.3. For this method, the effective focal length is now assumed to be a known parameter. Therefore, the values for G_y and G_z are determined during the calibration.

- Unknowns: $G_y, G_z, C_y, C_z, k_1, k_2, k_3, k_4, k_5, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R},$ and α_{23R} .
- Equations to solve: 5.1 (formed with 5.3 and 5.3), where

$$\begin{aligned} D_y(i) &= k_1 y_D(i) (y_D^2(i) + z_D^2(i)) + k_2 y_D^2(i) + k_3 z_D^2(i), \\ D_z(i) &= k_1 z_D(i) (y_D^2(i) + z_D^2(i)) + k_4 y_D^2(i) + k_5 z_D^2(i). \end{aligned}$$

2.3.9. *Special radial.* The “Special radial 1” model assumes that the radial curvature coefficients are different for both image plane coordinate axes. The mathematical equations, as implemented, can be found in Section 2.2.3. Single order distortion coefficients are used. The size and resolution of the camera image plane (G_y, G_z) is assumed known so that the effective focal length and the uncertainty scale factor S_y for G_y remain to be established.

- Unknowns: $f_e, S_y, C_y, C_z, k_1, k_2, X_{CTR}, Y_{CTR}, Z_{CTR}, \alpha_{12R}, \alpha_{13R},$ and α_{23R} .
- Equations to solve: 5.1 (formed with 5.3 and 5.3), where

$$\begin{aligned}
D_y(i) &= k_1 y_D(i) (y_D^2(i) + z_D^2(i)) , \\
D_z(i) &= k_2 z_D(i) (y_D^2(i) + z_D^2(i)) .
\end{aligned}$$

3. Conclusion

This section presented the calibration methods implemented and tested for this report. The next section contains some information about the techniques used to evaluate and rank all the different methods.

CHAPTER 6

Calibration performance evaluation

One of the first questions about any camera calibration is how accurately does it capture the image behavior? This information is used to estimate the performance or accuracy of the camera in the context of its use. Given the intrinsic calibration data, the measured coordinates of targets in the TCS (X_C , Y_C , Z_C) and the measured position of the point's image in the FCS (y_F , z_F), it is possible to define an error metric for the calibration technique in the Camera-space or in the Object-space.

This section contains the mathematical details of three different methods used to evaluate the performance of camera calibrations. The evaluation methods include the Camera-Space Error (*CSE*), the Object-Space Error (*OSE*) and the Relative Photosolution Error (*RPSE*). The section also introduces new equations, adapted for CSVS, to estimate the centroid error. *CSE* or *OSE* values significantly larger than the estimated centroid error indicate a lack of correlation between the calibration method and the measured data.

Two independent target arrays are used along with those evaluation methods. The first verification array consists of all the targets on the calibration board that have been relocated from the reference calibration position. The second array is an independent target configuration recreating the size and distance from the camera of a typical flight payload. This calibration verification setup is called the *volumetric array*.

1. Estimated centroid error

The methodology presented in [31] allows one to derive the estimated centroid error for a binary thresholded video signal as a function of the dot diameter. This information will

be used throughout this section to establish an error upper bound against which the various calibration evaluation methods will be compared. The purpose of the estimated centroid error is to put in perspective the magnitude of the calibration residuals when compared to the expected inherent error obtained during the sensing process.

Using similar triangles, the dot diameter in the image plane from the projection of a target located in the object-space is given by:

$$\begin{aligned} d_y &= \frac{D \cdot f_e}{R \cdot G_y}, \\ d_z &= \frac{D \cdot f_e}{R \cdot G_z}, \end{aligned}$$

where:

D : Real target diameter.

R : Distance between the camera optical center and a target in the Object-space.

d_y : Horizontal dot diameter.

d_z : Vertical dot diameter.

The estimated centroid error formulas, expressed as functions of the target image diameter in both the vertical and horizontal directions, are given by:

$$\begin{aligned} \delta_y &= \frac{0.027}{\sqrt{d_y}}, \\ \delta_z &= \frac{0.157}{\sqrt{d_z}}. \end{aligned}$$

The previous equations are more conservative than what is presented in [31]. The updated functions are the result of a special video dithering function within CSVS that improves its centroid measurement accuracy. Experimental and mathematical details about this subject are presented in [41].

2. Calibration evaluation on the calibration board

The calibration board used to perform the calibration verification tests is the same as the one presented in Figure 4.11. It consists of a flat, black painted surface (28.5 inches by

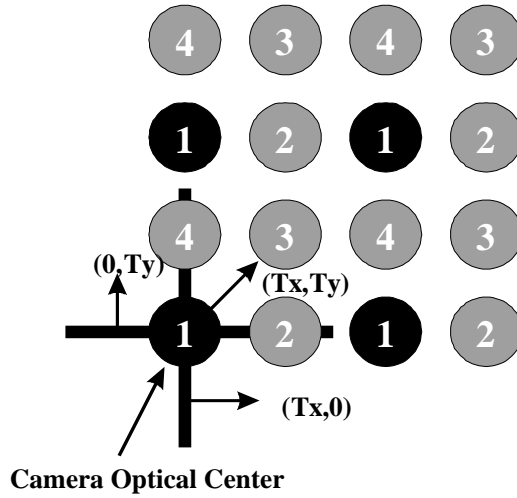
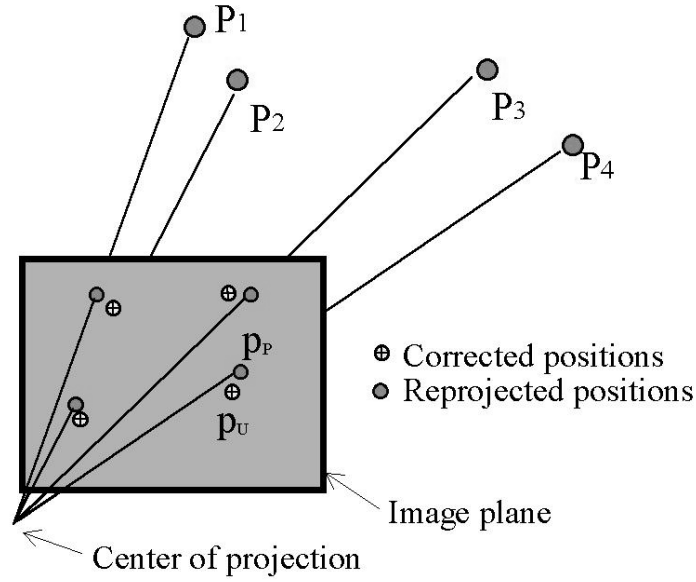


FIGURE 6.1. Dot locations for residuals calculation.

21.5 inches) with 165 white circular 0.62-inch dots, 1.75 inches apart and arranged in an 11 by 15 array.

In order to perform the calibration verification tests, the calibration board is first positioned perpendicular to the camera optical axis following the alignment procedures found in Section 1.1 of Chapter 5.

In order to obtain the most significant calibration verification results, it is important to ensure that the new centroids used to perform the tests are as far away as possible from the original calibration point. A practical methodology is illustrated in Figure 6.1. The dots labeled with 1 represent the original array position for the calibration board after the alignment routine is performed. Dots labeled with 2, 3, and 4 represent the new projected calibration targets after the camera (or the calibration board) has been translated parallel to the image plane by half the distance between two adjacent dots in the horizontal (T_y) and vertical (T_x) directions. At every location, the dot centroids can be recorded as part of the calibration verification process. It is expected however, that the most significant set of data comes from location number 3, where the dots are the farthest away from the original calibration dots.

FIGURE 6.2. Camera-space error (CSE).

For this study, the target centroids were recorded only at location 3 and therefore the centroids used for the evaluations were separate and independent from the centroids used to generate the camera calibrations.

2.1. Camera-Space Error (CSE) on the calibration board. One obvious error metric is the difference between the position of a measured and corrected centroid and the estimated true centroid from the camera photogrammetric solution. The estimated true centroid is determined by reprojecting the target array into the image plane. Figure 6.2 contains a graphical representation of the concept. The separation between the corrected target positions and the intersection of the photosolution reprojection lines with the image plane represents the Camera-Space Error (CSE).

The calculation of the CSE on the calibration board involves proceeding through the following steps:

1. The first step of the process is the positioning of the calibration board in the field-of-view of the camera as described in Section 2. This position is called T_E , for the evaluation position. There is one evaluation position per camera/lens, therefore all techniques are evaluated with identical measured centroids.

The target centroids in the FCS ($y_F(i)$, $z_F(i)$) are then recorded in the SVS. This gives 165 measured centroids. Each measured centroid consists of an average of 100 raw SVS centroids (1 per video frame) for each target. The SVS program records the centroids one target at a time following a raster scan pattern on the calibration board.

The camera calibration data can then be used to transpose the distorted centroids to their new corrected locations in the ICS ($y_U(i)$, $z_U(i)$). Equations found in Chapters 3 and 4 for the various calibration techniques are applied at this stage.

2. Using the target locations in the TCS, the corrected target image locations in the ICS and the camera intrinsic parameters, a photogrammetric solution is derived. The final vector defining the transformation from the CCS to the TCS using all the calibration board targets is called SCT_E .

The colinear relations of photogrammetry from Section 1, given in Equation 5.1 (formed with 5.3 and 5.3) are used to derive the mathematical expressions to be minimized by the LS fitting criterion.

The colinear conditions contain a total of six unknowns - three defining the position, as given by X_{CTE} , Y_{CTE} , Z_{CTE} , and three selected independent direction cosine terms (α_{12E} , α_{13E} , and α_{23E}) to define the orientation. The LS minimization algorithm developed for this study was implemented in Matlab and follows the principles mentioned in [5].

3. The reconstructed target location ($y_P(i)$, $z_P(i)$) is derived by projecting the point's 3D coordinates ($X_T(i)$, $Y_T(i)$, $Z_T(i)$) through the calibrated pinhole camera model. The formulas are similar to the ones given in Chapter 3.

$$\begin{aligned} y_p(i) &= \left(\frac{Y_{CTE} + X_T(i)\alpha_{21E} + Y_T(i)\alpha_{22E} + Z_T(i)\alpha_{23E}}{X_{CTE} + X_T(i)\alpha_{11E} + Y_T(i)\alpha_{12E} + Z_T(i)\alpha_{13E}} \right) \cdot f_e, \\ z_p(i) &= \left(\frac{Z_{CTE} + X_T(i)\alpha_{31E} + Y_T(i)\alpha_{32E} + Z_T(i)\alpha_{33E}}{X_{CTE} + X_T(i)\alpha_{11E} + Y_T(i)\alpha_{12E} + Z_T(i)\alpha_{13E}} \right) \cdot f_e. \end{aligned}$$

4. Using the data calculated in the previous stages, the formula for the Camera-Space Error (CSE) for all 165 points on the calibration board is given by:

$$(6.1) \quad CSE = \frac{1}{N} \sum_{i=1}^N \sqrt{(y_U(i) - y_P(i))^2 + (z_U(i) - z_P(i))^2}.$$

This last formula represents the quantity for the CSE used for the rest of the document.

Another solution to the consistency problem in the units is to normalize the CSE measurements, in millimeters, by the effective focal length. This idea was rejected since the video line unit is considered more intuitive than measuring error in terms of the ratio of the effective focal length.

5. The subsequent step is to derive the CSE upper error bound based on the estimated centroid error given in Section 1. Using the original calibration method data obtained with the different cameras, Table 6.1 presents the ideal calibration dot diameters for the calibration board.

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
D (in.)	0.62	0.62	0.62	0.62
f_e (mm)	13.5	12.4	6.0	8.0
R (in.)	35.7353	36.7091	25.2029	33.6468
G_y (mm/pixel)	0.013478	0.012115	0.008441	0.008439
G_z (mm/line)	0.015864	0.014221	0.009748	0.009791
d_y (pixels)	17.38	17.29	17.49	17.47
d_z (lines)	14.76	14.73	15.14	15.06

TABLE 6.1. Ideal calibration dot diameters.

From the image plane dot diameters just calculated, Table 6.2 presents the standard estimated centroid error.

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
δ_y (lines)	5.78E-02	5.78E-02	5.71E-02	5.72E-02
δ_z (pixels)	5.33E-02	5.34E-02	5.31E-02	5.31E-02
δ_z (lines)	4.52E-02	4.55E-02	4.60E-02	4.58E-02
$RMS\delta$ (lines)	7.34E-02	7.36E-02	7.33E-02	7.33E-02

TABLE 6.2. Estimated centroid error in the calibration board.

The theory presented in this section will be used in Chapter 7 to evaluate the calibration performance of the different cameras on the calibration board.

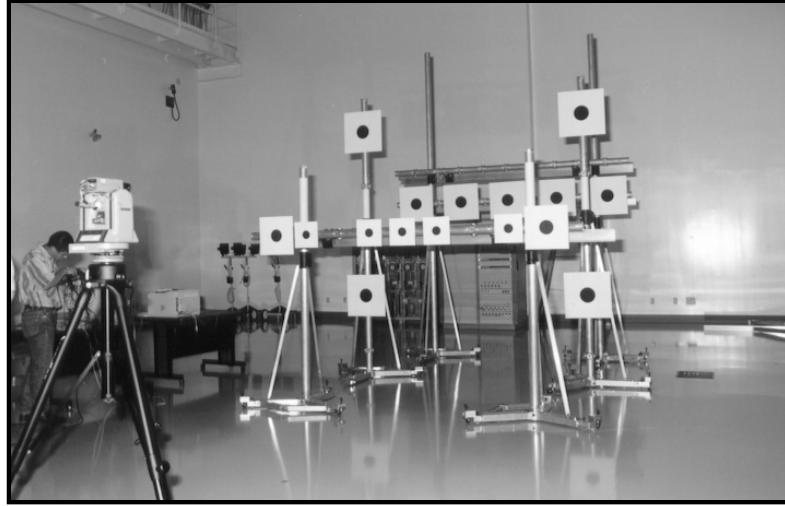


FIGURE 6.3. Example of a volumetric target array.

3. Calibration evaluation on the volumetric array

An effective way of measuring the proficiency of a given calibration is to use an independent test bench recreating the size and distance from the camera of a typical flight payload. This target array configuration is called the volumetric array and is the primary camera calibration evaluation setup used for this study. Figure 6.3 presents an example of a volumetric array, similar to the one used to generate the data for this thesis. Also displayed on the left-hand side, is the ground survey instrument used to measure the target locations relative to the TCS.

The volumetric array actually used for this study was designed with the same equipment shown in Figure 6.3 but with a different target configuration. Figure 6.4 provides a graphic representation of the relative target locations when viewed along the camera optical axis.

The volumetric array consists of a maximum of 18 targets subdivided into three different ranges (W, X, Y) in the volumetric array. The overall dimensions of the volumetric array are 10.64 feet deep by 14 feet wide by 9.4 feet high. These dimensions were chosen to match a standard camera 4:3 aspect ratio.

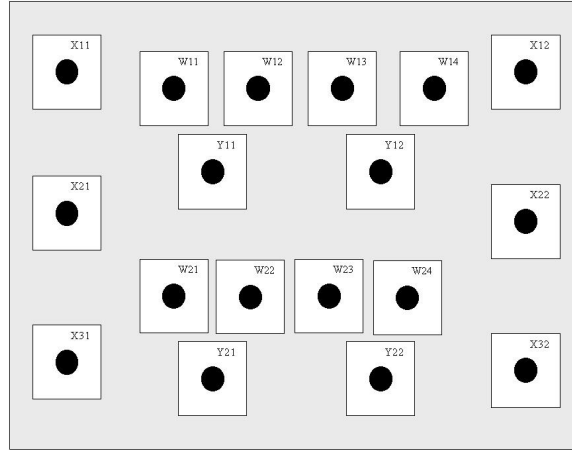


FIGURE 6.4. Volumetric target array configuration.

3.1. Camera-Space Error (CSE) on the volumetric array. The calculation of the CSE on the volumetric array uses the same mathematical principles previously developed for the CSE on the calibration board. The only difference comes from the approach used to position the targets in the camera field of view. The different steps involved in calculating the CSE on the volumetric array are now described:

1. The first step of the process is the positioning of the camera so that all the X-range targets from the volumetric array are projected at the boundaries of the field of view. Although the location of the volumetric array remains unchanged during the alignment process, its position is called T_E , for the evaluation position. There is one evaluation position per camera/lens therefore all techniques are evaluated with identical measured data. The position of the camera with respect to the volumetric array is different depending on the camera field of view, although the volumetric array appears almost the same in the camera image plane.

The target centroids in the FCS ($y_F(i)$, $z_F(i)$) are then recorded with the SVS. The number of target centroids might be less than 18 (maximum number of targets in the volumetric array) if the camera distance from the array causes some targets to be occluded by the array structure. Each measured centroid consists of an average of 1000 raw SVS centroids (1 per video frame). The data is collected eight targets at a time, until all the centroids of the visible targets in the field of view are recorded.

The camera calibration data can then be used to transpose the distorted centroids to their new corrected locations in the ICS ($y_U(i)$, $z_U(i)$). Equations found in Sections 3 and 4 for the various calibration techniques are applied at this stage.

2. Using the target locations in the TCS, the corrected target image locations in the ICS and the camera intrinsic parameters, a photogrammetric solution is derived. The final vector defining the transformation from the CCS to the TCS using all the calibration board targets is called SCT_E .

Steps similar to the ones presented in Section 2.1 are used to derive the CSE for all N targets forming the volumetric array.

3. The next step is to derive the CSE upper error bound based on the estimated centroid error given in Section 1.

Using an average target diameter of 7 inches, and using the range from the camera to the geometrical center of the volumetric array as measured by the different cameras, Table 6.3 presents the ideal calibration dot diameters.

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
D (in.)	7.0	7.0	7.0	7.0
f_e (mm)	13.5	12.4	6.0	8.0
R (in.)	291.686	292.966	242.375	287.294
G_y (mm/pixel)	0.013478	0.012115	0.008441	0.008439
G_z (mm/line)	0.015864	0.014221	0.009748	0.009791
d_y (pixels)	24.04	24.46	20.53	23.10
d_z (lines)	20.42	20.83	17.78	19.91

TABLE 6.3. Ideal dot diameters in the volumetric array.

From the image plane dot diameters just calculated, Table 6.4 presents the standard estimated centroid error.

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
δ_y (lines)	4.91E-02	4.86E-02	5.27E-02	4.98E-02
δ_z (pixels)	4.53E-02	4.49E-02	4.90E-02	4.62E-02
δ_z (lines)	3.85E-02	3.82E-02	4.24E-02	3.98E-02

TABLE 6.4. Estimated centroid errors in the volumetric array.

The theory presented in this section is used in Chapter 7 to evaluate the CSE in the volumetric array.

3.2. Object-Space Error (OSE) on the volumetric array. The inverse perspective projection problem can be used to measure the camera calibration accuracy in object-space. By projecting a corrected image point $(y_U(i), z_U(i))$ back through the camera model, it is possible to calculate the closest distance of approach between the image point's line-of-sight and the point in 3D object-space $(X_C(i), Y_C(i), Z_C(i))$ that was supposed to have cast the image. The 3-D error measurement is called the Object-Space Error (OSE). This new calibration performance metric is best suited for arrays with sufficient depth such as the volumetric array and was therefore omitted for the data on the calibration board. The different steps involved in calculating the OSE on the volumetric array are now described:

1. The data collected during steps 1 and 2 from the previous section is used to derive the OSE .
2. Using the data calculated in the previous stages, the formula for the OSE for all targets in the volumetric array is given by:

$$(6.2) \quad OSE = \frac{1}{N} \sum_{i=1}^N \sqrt{(X_C(i) - f_e \cdot t(i))^2 + (Y_C(i) - y_U(i) \cdot t(i))^2 + (Z_C(i) - z_U(i) \cdot t(i))^2},$$

where

$$t(i) = \frac{X_C(i) \cdot f_e + Y_C(i) \cdot y_U(i) + Z_C(i) \cdot z_U(i)}{f_e^2 + y_U(i)^2 + z_U(i)^2}.$$

In the previous equations, N represents the number of targets in the volumetric array, f_e represents the camera effective focal length, and the vector $(y_U(i), z_U(i))$ forms the corrected target coordinates. The term $t(i)$ is the scale factor used to convert the camera-space measurements into the object-space measurements.

In the ideal situation the OSE value would be zero, indicating a perfect fit between the image point's line-of-sight and the point in 3D object-space.

3. Theoretically, it is possible to determine the range of theoretical values for the OSE considering the estimated centroid errors for the targets located in the center and at the corners (furthest distance from the center) of the volumetric array. Once

again, the standard deviations for the dot locations are calculated using the theory in Section 1.

The formula used to calculate OSE on the volumetric array, taking into consideration the estimated centroid errors, is given by:

$$(6.3) \quad OSE = \frac{1}{N} \sum_{i=1}^N \sqrt{(X_C(i) - f_e \cdot t(i))^2 + (Y_C(i) - (y_U(i) + \delta_y) \cdot t(i))^2 + (Z_C(i) - (z_U(i) + \delta_z) \cdot t(i))^2},$$

where

$$t(i) = \frac{X_C(i) \cdot f_e + Y_C(i) \cdot (y_U(i) + \delta_y) + Z_C(i) \cdot (z_U(i) + \delta_z)}{f_e^2 + (y_U(i) + \delta_y)^2 + (z_U(i) + \delta_z)^2}.$$

In the previous equations, the terms $(y_U(i) + \delta_y)$ and $(z_U(i) + \delta_z)$ are the unification of the ideal dot location and the estimated centroid error.

Table 6.5 presents a compilation of the theoretical OSE for the different cameras for targets at the center of the field of view.

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
<i>Range</i> (in.)	291.686	292.966	242.375	287.294
δ_y (mm)	6.62E-04	5.89E-04	4.44E-04	4.20E-04
δ_z (mm)	7.18E-04	6.38E-04	4.78E-04	4.52E-04
$t(i)$ (in./mm)	21.606	23.626	40.396	35.912
OSE (in.)	2.11E-02	2.05E-02	2.64E-02	2.22E-02

TABLE 6.5. OSE in the volumetric array at the center of the field of view.

Table 6.6 presents a compilation of the theoretical OSE for the different cameras for a target which is located at the bottom right corner of the field of view.

The theory presented in this section is used in Chapter 7 to evaluate the OSE in the volumetric array.

3.3. Relative photosolution error ($RPSE$) on the volumetric array. The last performance evaluation test conducted on the volumetric array is called Relative photosolution. The calibration errors are measured from two independent full six degree-of-freedom

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
<i>Range</i> (in.)	291.686	292.966	242.375	287.294
Y_T (in.)	84.025	84.025	84.025	84.025
Z_T (in.)	48.113	48.113	48.113	48.113
y_U (mm)	3.889	3.556	2.080	2.340
z_U (mm)	2.227	2.036	1.191	1.340
δ_y (mm)	6.62E-04	5.89E-04	4.44E-04	4.20E-04
δ_z (mm)	7.18E-04	6.38E-04	4.78E-04	4.52E-04
$t(i)$ (in./mm)	21.606	23.626	40.394	35.911
<i>OSE</i> (in.)	2.01E-02	1.96E-02	2.46E-02	2.11E-02

TABLE 6.6. *OSE* for target at the extremity of the field-of-view.

photosolutions (SCT_1 and SCT_2) on various sub-arrays defined to compare different regions of the image plane. Assuming that the target locations in the TCS ($X_T(i)$, $Y_T(i)$, $Z_T(i)$) from both arrays are defined with respect to the same location in object-space, the values for both photosolutions (SCT_1 and SCT_2) should be the same.

For the purpose of improving the photosolution convergence speed, an intermediate vector (SPT) was used to define the transformation between the geometric center of individual arrays and the common survey point for the entire volumetric array. Stated differently, two TCS are defined and are located at the geometric center of each target array, i.e. at the center of mass for all the targets. The two TCS are related to a common point P through the SP_1T_1 and SP_2T_2 vectors. The total chain of transformation used to measure the level of relative errors in SCT_1 and SCT_2 is given as follows:

$$(6.4) \quad SP_1P_2 = SP_1T_1 \cdot (SCT_1)^{-1} \cdot SCT_2 \cdot (SP_2T_2)^{-1}.$$

The vector SP_1P_2 represents the relative photosolution error ($RPSE$) as measured by the gap from the two arrays from the camera to the common point in the volumetric array. Therefore, it is expected that in the ideal case SP_1P_2 should be zero for all array pairs, i.e. $P_1 = P_2$.

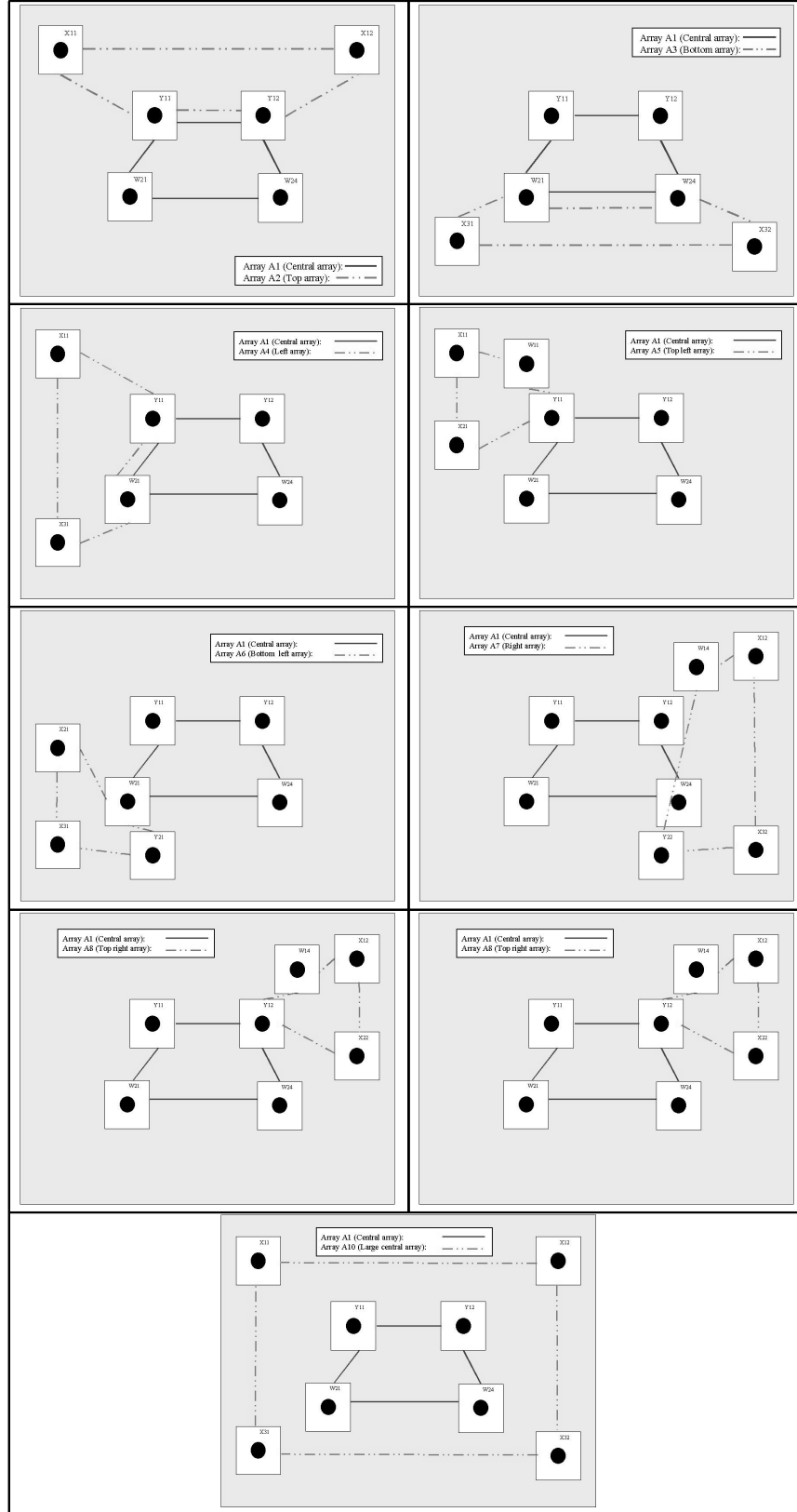
The LS minimization and the matrix transformation algorithms developed to evaluate the $RPSE$ were implemented in Matlab. The different steps involved in calculating the $RPSE$ on the volumetric array are now described:

1. As mentioned before, the volumetric array is composed of a maximum of 18 circular surveyed targets, whose projected image centroids are recorded using the SVS. The measured corrected centroids obtained from steps 1 and 2 in Section 2.1 are used again for this test.
2. The second step involves grouping different targets to form the different arrays. The intent of the grouping is to form arrays located in different regions of the camera field of view. It is typically observed that the center 25% of the image plane has the least distortions and therefore a central array is used as a reference array (P_1). Figure 6.5 presents the different array configurations used during the evaluation of the different camera calibration techniques.
3. Using the target locations in the TCS from both arrays, the corrected target image locations in the ICS and the camera intrinsic parameters, two photogrammetric solutions are derived (SCT_{E1} and SCT_{E2}). The final vector defining the transformation from the CCS to the TCS using all the calibration board targets is called SCT_E . The colinear relations of photogrammetry from Chapter 3 are used to obtain SCT_{E1} and SCT_{E2} .
4. From the two photosolutions (SCT_{E1} and SCT_{E2}) and the SPT_1 and SPT_2 vectors, the relative photosolution error ($RPSE$) is derived.

The theory presented in this section is used in Chapter 7 to evaluate the $RPSE$ in the volumetric array.

4. Conclusion

All the calibration verification methods mentioned above are used in this thesis to measure the performance of the different camera calibration methods. The next section contains the test results, procedures and analysis.

FIGURE 6.5. *RPSE* arrays definition.

CHAPTER 7

Test results

This Chapter contains the results of the comparative performance analysis of the camera calibration methods described in Chapter 5. As a reminder, three types of measurements were used to assess the performance of the calibration techniques:

1. *Camera-Space Error (CSE)*: describes the centroid measurement residuals on the camera image plane.
2. *Object-Space Error (OSE)*: represents the target location measurement residuals in 3D space.
3. *Relative Photosolution Error (RPSE)*: measures the error in the relative transformation between two arrays in 3D space.

The *CSE* was measured on both the calibration board and the volumetric array, however the *OSE* and *RPSE* were only measured for the volumetric array. Details about the different evaluation techniques, together with testing procedures, can be found in Chapter 6.

The various calibration methods are judged through their respective averaged residual errors which, assuming normal distributions, are grouped and ranked using a Student's *t*-test criteria. Groups are formed among candidates which have mean error values statistically similar to each other. Normally, a *t*-test significance value below 95% is the statistically accepted limit for grouping. A compilation of the test data appears in five tables which ranks the methods initially using an average of the group ranking, followed by the average residual errors. In our opinion, the consistency of a method should have precedence over the overall measured residual errors.

For details about the Student’s t-test used, or to review the complete test results of each individual camera, please refer to Appendix C.

1. Results of the *CSE* tests on the calibration board

Table 7.1 contains a compilation of the *CSE* test results.

Order	Methods	Average ranking	Average error (lines)
1	Lookup compensation method	1.00	0.055
2	Polynomial smoothing method with Lookup compensation method	1.50	0.064
3	Averaging of multiple maps	2.00	0.072
Standard centroid error			0.073
4	Original method	2.25	0.080
5	Polynomial smoothing method	2.75	0.088
6	Radial - fixed gain ratios	3.25	0.199
7	Radial - fixed focal length	3.25	0.207
8	Radial and decentering - 2 steps	3.50	0.105
9	Radial and second order	3.50	0.223
10	Radial and decentering	3.50	0.225
11	Special radial	3.50	0.229
12	Radial and thin prism	3.50	0.230
13	Radial - different data sets	3.75	0.230
14	Radial and tangential	4.00	0.114

TABLE 7.1. Calibration performance from the *CSE* on the calibration board.

The *CSE* obtained using the calibration board targets gives an indication of the level of distortion correction. The great number of sample points allows an unambiguous grouping of the methods based on their measured image plane residuals.

An overall review of Table 7.1 shows that all generic calibration methods, more specifically methods using lookup compensation, outperformed the parametric methods. Polynomial smoothing, used alone, did not improve nor really degrade the performance of the original method.

Parametric methods using simpler lens distortion models better captured the overall system’s aberrations. The “radial and tangential” and “radial and decentering - 2 steps” come out with relatively decent performances since calibration parameters could not be obtained for the ITVC due to the non convergence of our calibration method.

The data also indicates that three generic methods obtained average results that were better than the value for the estimated centroid error.

2. Results of the *CSE* tests on the volumetric array

Table 7.2 presents the results of the *CSE* tests performed on the volumetric array.

Order	Methods	Average ranking	Average error (lines)
Standard centroid error			0.064
1	Lookup compensation method	1.25	0.334
2	Original method	1.25	0.335
3	Radial and decentering	1.25	0.378
4	Radial and tangential	1.50	0.303
5	Polynomial smoothing method with Lookup compensation method	1.50	0.338
6	Polynomial smoothing method	1.50	0.340
7	Averaging of multiple maps	1.50	0.347
8	Radial and second order	1.50	0.408
9	Special radial	1.50	0.410
10	Radial - different data sets	1.50	0.416
11	Radial and thin prism	1.50	0.418
12	Radial - fixed gain ratios	1.50	0.423
13	Radial and decentering - 2 steps	1.75	0.328
14	Radial - fixed focal length	1.75	0.441

TABLE 7.2. Calibration performance from the *CSE* on the volumetric array.

The *CSE* obtained from fitting the photogrammetric equations on all the visible volumetric targets may provide insight with the quality of the different calibration methods. Since fewer sample points were taken, the task of separating the methods became more delicate.

Table 7.2 indicates that, on average, the generic calibration methods, specifically the “lookup compensation” and the “original” methods, outperform the parametric methods. The “radial and tangential” and “radial and decentering” methods obtain the best results for the parametric methods.

The results also show that all methods perform approximately the same when compared to the magnitude of the standard centroid error.

3. Results of the *OSE* tests on the volumetric array

Table 7.3 presents the results of the *OSE* tests performed on the volumetric array.

Order	Methods	Average ranking	Average error (inches)
Standard centroid error			0.021
1	Lookup compensation method	1.00	0.095
2	Original method	1.00	0.096
3	Polynomial smoothing method with Lookup compensation method	1.00	0.096
4	Polynomial smoothing method	1.00	0.097
5	Averaging of multiple maps	1.00	0.100
6	Radial and second order	1.25	0.116
7	Special radial	1.25	0.117
8	Radial - different data sets	1.25	0.119
9	Radial and thin prism	1.25	0.119
10	Radial and decentering	1.25	0.119
11	Radial - fixed gain ratios	1.25	0.120
12	Radial and tangential	1.50	0.089
13	Radial and decentering - 2 steps	1.50	0.096
14	Radial - fixed focal length	1.50	0.126

TABLE 7.3. Calibration performance from the *OSE* on the volumetric array.

The data from Table 7.3 agrees with the *CSE* data from Section 1. That is, all generic calibration methods outperform the parametric methods. Once again, the results also show that all methods perform approximately the same when compared with the magnitude and the standard centroid error.

4. Results of the *RPSE* tests on the volumetric array

Tables 7.4 and 7.5 summarize the results of the *RPSE* tests performed on the volumetric array.

The six degree-of-freedom relative solution derived from the volumetric array is a method by which the global accuracy of the camera calibration can be measured.

Tables 7.4 and 7.5 reveal that similar results were obtained from both categories, the parametric and the generic methods producing slightly better results on the translation and the rotation errors, respectively.

Order	Methods	Average ranking	Average error (inches)
1	Radial - fixed focal length	1.00	0.494
2	Special radial	1.00	0.571
3	Radial - different data sets	1.25	0.584
4	Polynomial smoothing method	1.25	0.601
5	Radial and second order	1.25	0.605
6	Radial and thin prism	1.25	0.606
7	Radial and decentering	1.25	0.608
8	Radial - fixed gain ratios	1.25	0.610
9	Polynomial smoothing method with Lookup compensation method	1.25	0.613
10	Original method	1.25	0.655
11	Averaging of multiple maps	1.25	0.657
12	Lookup compensation method	1.25	0.670
13	Radial and tangential	1.50	0.516
14	Radial and decentering - 2 steps	1.50	0.526

TABLE 7.4. Calibration performance from the *RPSE* (translation) on the volumetric array.

Order	Methods	Average ranking	Average error (degrees)
1	Polynomial smoothing method with Lookup compensation method	1.00	0.116
2	Polynomial smoothing method	1.00	0.118
3	Lookup compensation method	1.00	0.118
4	Original method	1.00	0.120
5	Radial - fixed focal length	1.00	0.122
6	Averaging of multiple maps	1.00	0.130
7	Radial and second order	1.25	0.157
8	Radial - different data sets	1.25	0.161
9	Radial and decentering	1.25	0.162
10	Special radial	1.25	0.163
11	Radial and thin prism	1.25	0.165
12	Radial - fixed gain ratios	1.25	0.166
13	Radial and tangential	1.50	0.124
14	Radial and decentering - 2 steps	1.50	0.132

TABLE 7.5. Calibration performance from the *RPSE* (rotation) on the volumetric array.

5. Performance of the different calibration methods

The four camera configurations tested provided a good sample of foreseen possible cases for CSVS; fixed and variable focus lenses, CCD and ICCD technologies, with and without

electronic signal enhancement, etc. The results obtained did not reveal that one particular camera calibration technique stood out from the rest as being better by a considerable margin. The generic method performed as well as, and often better than, the parametric methods for all cameras. This conclusion is drawn from the data presented in Sections 1 to 3. It was observed that the *OSE* did not supply additional insight into the performance of the camera calibration that the *CSE* had not already provided.

The Tables 7.6 and 7.7 contain a summary of the *CSE* collected on the calibration board and on the volumetric array. The last two columns are the *CSE* normalized with respect to the estimated centroid error (δ).

	Generic (lines)	Parametric (lines)	Centroid err. (lines)	Generic (normalized)	Parametric (normalized)
CTVC	0.054	0.060	0.073	0.737	0.822
ITVC	0.090	0.548	0.074	1.219	7.447
Cohu (6 mm)	0.089	0.187	0.073	1.212	2.549
Cohu (8 mm)	0.054	0.086	0.073	0.735	1.168
Average	0.072	0.220	0.073	0.976	2.996

TABLE 7.6. Summary of *CSE* results on the calibration board.

	Generic (lines)	Parametric (lines)	Centroid err. (lines)	Generic (normalized)	Parametric (normalized)
CTVC	0.444	0.424	0.062	7.112	6.791
ITVC	0.399	0.682	0.062	6.458	11.032
Cohu (6 mm)	0.260	0.274	0.042	6.137	6.459
Cohu (8 mm)	0.252	0.269	0.064	3.953	4.216
Average	0.339	0.412	0.058	5.915	7.125

TABLE 7.7. Summary of *CSE* results on the volumetric array.

The Tables 7.6 and 7.7 indicate that the performance of all calibration methods worsen when going from the calibration board to the volumetric array and that the error is more important on the ITVC.

The next subsection presents a review of the different methods studied from the perspective of the results obtained in this chapter.

5.1. Generic methods review. Despite their apparent simplicity, the generic methods proved to be very robust i.e. would always generate calibration data even with the

severe distortions seen for the Cohu with a 6 mm lens and the ITVC. Also, the accuracy measured was always comparable to, or clearly better than, the more complex parametric methods. One particularly interesting aspect of the generic methods is the ability to deal with special non-uniform distortions that are very hard to model mathematically.

On the other hand, as mentioned earlier, the generic calibration methods accuracy is strongly dependent on the quality of the alignment between the calibration board and the optical axis.

The result of misalignment error on the generic methods calibration data is twofold. First, a correction map is generated assuming an artificial pitch or yaw attitude (also inducing X , Y and Z errors from degree-of-freedom coupling) that will be apparent during the real operation or during the volumetric testing. Second, the intrinsic camera parameters, especially the gain ratios, will be biased in the direction of the slant. As presented in the next equation, it is good practice to calculate the calibrated image size to confirm that the 4:3 standard NTSC aspect ratio is preserved.

$$(7.1) \quad \frac{G_y N_p}{G_z N_l} = 1.3$$

In Equation 7.1, N_p and N_l represent the number of pixels in a line and the number of lines in a frame defined by the frame grabber A/D converter. Table 7.8 contains the results from the gain ratio tests from the cameras tested.

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
G_y (mm/pixel)	0.013478	0.012115	0.008441	0.008439
G_z (mm/line)	0.015864	0.014221	0.009748	0.009791
N_p (pixels)	752	752	752	752
N_l (lines)	480	480	480	480
Vert. size (mm)	10.135	9.110	6.348	6.346
Horiz. size (mm)	7.615	6.826	4.679	4.700
Ratio	1.331	1.335	1.357	1.350
Error (%)	0.172	-0.099	-1.746	-1.275

TABLE 7.8. Generic methods gain ratio test results.

As explained in [55] errors of up to $\pm 1.5\%$ are considered typical, but in cases where lenses with a high degree of curvature (short focal lengths) are used, higher numbers are expected (Cohu with a 6 mm lens). The additional error is conceivably caused by the fact

that the gain ratios are calculated using a small subset of the calibration points at only two locations on the X-motion rail (ref. Chapter 5, Section 1.1).

The data presented in this chapter indicates that lookup compensation methods enhance the performance of the original method. Other calibration improvements, such as averaging multiple maps or polynomial smoothing do not necessarily have a big impact on the final results. It was observed that the map averaging technique tends to worsen the correction map when long extrapolations are used.

5.2. Parametric methods review. The modeling of distortion patterns is intuitively a powerful approach to the camera calibration problem. Simple and wide-spread theories about camera optics are correlated with the sensor data to produce a succinct list of calibrated parameters. In addition, a rigorous and lengthy alignment process is avoided when recreating a three-dimensional target array by moving the calibration board by known increments. The calibration board is aligned perpendicular to the X-motion rail and is not necessarily aligned with the camera. The only requirement is to position the camera so that the calibration board fills most of the field of view. The majority of the work is therefore spent waiting for the vision system to record the multitude of target dots and moving the calibration board to various positions.

Contrary to the generic methods, all the intrinsic and extrinsic parameters are calculated at the same time, making all the parameters “optimal” in a least-squares sense. Table 7.9 contains the results from the gain ratio tests from the cameras tested.

	CTVC	ITVC	COHU (6 mm)	COHU (8 mm)
G_y (mm/pixel)	0.01364	0.01219	0.00817	0.00829
G_z (mm/line)	0.01600	0.01428	0.00953	0.00967
N_p (pixels)	752	752	752	752
N_l (lines)	480	480	480	480
Vert. size (mm)	10.257	9.167	6.144	6.234
Horiz. size (mm)	7.680	6.854	4.574	4.642
Ratio	1.336	1.337	1.343	1.343
Error (%)	-0.169	-0.303	-0.732	-0.732

TABLE 7.9. Parametric methods gain ratio test results.

Full accuracy is also obtained when correcting for lens distortions by directly calculating the corrections from mathematical equations, hence avoiding the bi-linear interpolation

technique used with the correction map method. As presented before, the bi-linear interpolation violates the fact that the distortions are typically growing from the optical axis following a high order polynomial pattern and also that the reference correction points are not the ones measured during the calibration.

In spite of those advantages, the parametric methods have some substantial limitations for a product like the CSVS. First, the system is operated with cameras and lenses engineered in such a way that their characteristics depart from the models presented in Chapter 2. The best example is the ITVC CCD and Light Intensifier, which exhibits uncommon scene distortions difficult to model. Furthermore, complex lenses like the Fujinon 15X or Angénieux 8X (see Appendix A) are designed with an impressive number of glass elements such that the non-radial distortions from element to element are self-correcting leaving a small probability that a simple parametric model could effectively correct for the leftover aberrations. Hence, the global accuracy of the parametric models is limited by the correlation between the distortion model and the true nature of the camera distortion.

When the distortion model fits the data with difficulty, the rest of the parameters are forced to depart from their true value to obtain the best-fit solution in a least-squares fashion. As a result, when substantial calibration accuracy is required, many different models have to be tested and careful analysis has to be performed to select the best calibration. However, fewer sets of models have to be tested if the lens characterization performed by the manufacturer is available. This data is typically provided for high-quality products.

At the defense of the parametric techniques, better parametric calibrations might have been achieved by using calibration board positions that filled the field of view at all times. It was discovered after plotting the calibration centroids, that moving the calibration board away from the camera left only a few calibration points outside the 70% field of view area.

Lastly, in the cases studied we found that better calibration performance was achieved when fewer distortion coefficients were used. Adding thin prism and decentering effects was occasionally a benefit, but more often degraded the global accuracy. As a final recommendation, tests should be done to verify whether certain variables take on a presumed value under a selected significance level $(1 - \alpha)$, based on the estimated variance of the variable. Such tests are presented in [62].

6. General use of the calibration evaluation methods

The approach adopted to evaluate the various calibration methods was first to measure the error level or calibration residuals on the calibration board, and second to repeat the tests with an independent system such as the volumetric array.

The *CSE* on the calibration board is mainly used to verify the ability of the calibration method to correct for distortions. The performance threshold can be established using the estimated centroid error presented in Chapter 6.

The *CSE*, *OSE*, and *RPSE* tests performed on the independent volumetric array present a global assessment of the quality of the calibration. All intrinsic camera parameters contribute to the error estimation process. For the *CSE*, *OSE*, a performance criteria can be derived using an extension of the estimated centroid error. In the ideal case, the *RPSE* should be equal to zero.

Comparing the results from the tests on both the calibration board and the volumetric array indicates that the calibration board testing is not enough to validate the calibration. Analyzing only the calibration board error data would have made the task of predicting the large relative array solution errors almost impossible.

It is believed that the tolerances used to design and assemble our calibration rig might explain the difference in accuracy obtained between the calibration board and the volumetric array tests.

7. Conclusion

This chapter presented the camera calibration performance results obtained from the different methods. The next chapter completes this thesis by presenting an overview of the topics discussed and the results obtained. The very last section deals with some unanswered camera calibration questions and summarizes new research areas that we would like to pursue.

CHAPTER 8

Conclusion

This thesis presents a review of various techniques and concepts used in the field of camera calibration. Camera calibration is the art of mathematically seeking the “intrinsic” and “extrinsic” parameters, i.e. the information that relates to the camera optics (including the lens distortions) and the information that describes the location and orientation of the camera in 3D space.

The calibration techniques studied can be classified into two distinct categories: the “generic” and the “parametric” calibration techniques. Contrary to the parametric approach, the generic calibration methods do not assume any distortion equation to fit the distorted data. Rather, a correction map for the entire image plane is generated.

The performance of the different calibrations was evaluated at two levels. The first level verifies the validity and accuracy of the distortion correction. The task is performed using the calibration apparatus, recording additional data different from the original calibration information. A second level of testing uses a volumetric array of size similar to the real operation environment to verify the accuracy of the general camera parameters.

The analysis results gathered for this report were compared with benchmark values obtained from the estimated centroid error (δ) presented in Section 1 of Chapter 6.

The following subsections present a summary of salient topics and test results intended to point towards new directions to improve the accuracy and repeatability of the camera calibration for the CSVS.

1. Preferred calibration method

Tables 8.1 and 8.2 show a summary of the Camera-Space Errors collected on the calibration board and on the volumetric array. The last two columns are the *CSEs* normalized with respect to the estimated centroid error (δ).

	Generic (lines)	Parametric (lines)	Centroid err. (lines)	Generic (normalized)	Parametric (normalized)
CTVC	0.054	0.060	0.073	0.737	0.822
ITVC	0.090	0.548	0.074	1.219	7.447
Cohu (6 mm)	0.089	0.187	0.073	1.212	2.549
Cohu (8 mm)	0.054	0.086	0.073	0.735	1.168
Average	0.072	0.220	0.073	0.976	2.996

TABLE 8.1. Summary of *CSE* results on the calibration board.

	Generic (lines)	Parametric (lines)	Centroid err. (lines)	Generic (normalized)	Parametric (normalized)
CTVC	0.444	0.424	0.062	7.112	6.791
ITVC	0.399	0.682	0.062	6.458	11.032
Cohu (6 mm)	0.260	0.274	0.042	6.137	6.459
Cohu (8 mm)	0.252	0.269	0.064	3.953	4.216
Average	0.339	0.412	0.058	5.915	7.125

TABLE 8.2. Summary of *CSE* results on the volumetric array.

The calibration verification data indicates that the generic methods produced comparable or clearly better results than the parametric methods. Furthermore, a definite improvement is observed when lookup compensation techniques are used.

2. Camera configuration

The applicability of a given set of camera parameters is more or less limited to the camera configuration and the lighting conditions that were selected at the time of the calibration. Investigating the camera response to changing conditions should be given a fairly high priority since the effects could be quite pronounced on the photosolution accuracy.

3. Calibration rig

The calibration rig is the fundamental component of the calibration task. The level of tolerance and functionality specified to build the apparatus has a direct effect on the

accuracy of the camera calibration task. It serves as “ground truth” for establishing the projective scaling obtained with a given camera. Special attention to construction tolerances is necessary to ensure that physical errors are below a level that can be detected by the CCD camera. Also, the markings on the calibration board should be planned to simultaneously maximize the target size and minimize the interpolation distance between the reference targets.

4. Camera calibration evaluation

The use of a volumetric array proved to be useful to uncover the weaknesses of the calibration rig created specifically for this study. However, as the calibration rig becomes more and more dependable, the goal is to develop means of fully verifying the calibrations using only the calibration board. Once equivalent evaluation results are obtained on the volumetric array as on the calibration board, the volumetric array will be dispensed.

Ultimately, to decide if the level of calibration is sufficient, it is desirable to determine the effect of the calibration errors in the context of an operation. The calibration procedure should determine error estimates for the intrinsic and extrinsic parameters. The error estimates would serve two purposes. First, the error estimates would form part of the criteria for deciding if the calibration is successful. Second, these error estimates would be required by error modeling tools such as the SVS Accuracy Model (SAM) in order to predict the final six degree-of-freedom errors at the payload mating interface.

For the focal length, gain ratios, optical center, and the position and orientation of the camera coordinate system, the standard deviation of error should be provided. The camera/lens residual error is modeled by SAM as a constant error, plus an error that is proportional to the square of the distance of the target image from the optical center. SAM requires a constant and a quadratic error coefficient for horizontal and vertical errors in the image plane, for a total of four coefficients.

5. Areas of future research

More testing will be necessary to determine the exact cause of the apparent change in camera calibration performance between the calibration board and the volumetric array. Therefore, most generic methods will be revisited, the existing calibration rig will be verified

and improved, and the volumetric array will be re-surveyed. Following this, a new attempt will be made to obtain results of similar accuracy between the two arrays.

Work to come will include the development of a series of tests intended to measure the range of applicability of a given set of calibration parameters. This effort will focus on measuring, and then expressing mathematically, the centroid bias from nominal operation conditions, as lighting conditions vary.

Future areas of research will include the development of new methods for camera self-calibrations with a minimum of *a priori* information. We believe that this field of research will become increasingly popular as vision systems for remote sensing operations rely more heavily on CCD cameras equipped with zoom lenses.

GLOSSARY

A/C, ADC	Analog to Digital Conversion
CBA	Calibration Board Assembly
CLCUTIL	Camera Lens Calibration Utility
CSLC	Cubic Spline Smoothing with Lookup Compensation
CCD	Charged Coupled Device
CCE	Charge Conversion Efficiency
CCS	Camera Coordinate System
CMA	Camera Mounting Apparatus
CSA	Canadian Space Agency
CSE	Camera-Space Error
DOF	Degrees of freedom
FCS	Frame Coordinate System
HCS	Housing Coordinate System
HFOV	Horizontal Field of View
HPS	High-order Polynomial Smoothing
ICCD	Intensified Charged Coupled Device
ICS	Image-plane Coordinate System
LS	Least-Squares
LSA	Longitudinal Spherical Aberration
MTF	Modulation Transfer Function
NASA	National Aeronautics and Space Administration
NRCC	National Research Council of Canada
OSE	Object-Space Error
PC	Personal Computer

PS	Polynomial Smoothing
PSLC	Polynomial Smoothing with Lookup Compensation
PTU	Pan-Tilt Unit
RLS	Ring Light Source
RMS	Remote Manipulator System
RPSE	Relative Photosolution Error
SAM	SVS Accuracy Model
SVS	Space Vision System
SVU	Space Vision Unit
STS	Space Transportation System
TCS	Target Coordinate System
TSA	Transversal Spherical Aberration
VSCL	Vision System Certification Laboratory

REFERENCES

- [1] J. Aloimonos and A. Bandyopadhyay. Active vision. In *First International Conference on Computer Vision, London*, pages 35–54, June 1987.
- [2] American Society of Photogrammetry. *Manual of Photogrammetry*, 1944,1952,1966,1980.
- [3] ANSI/SMPTE. *Composite Analog Video Signal - NTSC for Studio Applications*, 1994.
- [4] R. Bajcsy. Active perception. *Proceedings of the IEEE*, 76(8):996–1005, August 1988.
- [5] G.L. Basso and R. B. Kulchyski. Least-squares photogrammetry algorithm formulation. Technical report, National Research Council Canada, 1986.
- [6] P. Beardsley, D. Murray, and A. Zisserman. Camera calibration using multiple images. *European Conference on Computer Vision*, pages 312–320, 1992.
- [7] H.A. Beyer. Accurate calibration of CCD cameras. In *Conference on Computer Vision and Pattern Recognition*, pages 96–101, 1992.
- [8] R.A. Boie and I.J. Cox. An analysis of camera noise. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 14(6):671–674, June 1992.
- [9] P. Brand. *Reconstruction tridimensionnelle à partir d’une caméra en mouvement: de l’influence de la précision*. PhD thesis, Universite Claude Bernard, Lyon I, October 1995.
- [10] P. Brand, R. Mohr, and P. Bobet. Distorsions optiques: Correction dans un modèle projectif. Technical Report Rapport technique no. 1933, Unit de Recherche INRIA Rocquencourt, Juin 1993.

- [11] D. C. Brown. Decentering distortion of lenses. *Photogrammetric engineering and remote sensing*, pages 444–462, May 1996.
- [12] D.A. Butler and P.K. Pierson. A distortion-correction scheme for industrial machine-vision applications. *IEEE Transactions on Robotics and Automation*, 7(4):546–551, August 1991.
- [13] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(6), November 1986.
- [14] C. Cauchois, D. Meizel, and E. Brassard. Calibration du capteur de vision omnidirectionnelle: SYCLOP. Technical report, Groupe de Recherche sur l’Analyse et la Commande des Systèmes (GRACSY), Août 1998.
- [15] G. Champleboux, S. Lavallée, and P. Cinquin. Accurate calibration of cameras and range imaging sensors: the NPBS method. In *IEEE International Conference on Robotics and Automation, Nice, France*, May 1992.
- [16] J. J. Craig. *Introduction to Robotics, Mechanics and Control*. Addison-Wesley Publishing, 1992.
- [17] M. Devy, V. Garric, and J. J. Orteu. Camera calibration from multiple views of a 2D object, using a global non-linear minimization method. In *IEEE, International Conference on Intelligent Robots and Systems*, pages 1583–1589, 1997.
- [18] J. H. Dickson. *Optical Instruments and Techniques*. Oriel Press, 1969.
- [19] W. Faig. Calibration of close-range photogrammetric systems: Mathematical formulation. *Photogrammetric Engineering and Remote Sensing*, 41(12):1479–1486, December 1975.
- [20] O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.
- [21] O. Faugeras, Q.-T. Luong, and S.J. Maybank. Camera self-calibration: Theory and experiments. In *European Conference on Computer Vision*, 1992.
- [22] O. Faugeras and G. Toscani. Camera calibration for 3D computer vision. In *International Workshop on Industrial Applications of Machine Vision and Machine Intelligence, Seiken Symposium, Tokyo, Japan*, February 1987.

- [23] F.G. Fryer and C.D. Brown. Lens distortion for close range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 52(1):51–58, January 1986.
- [24] J. M. Geary. *Introduction to Optical Testing*. USA: SPIE Optical Engineering Press, 1993.
- [25] J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 1968.
- [26] A. Goshtasby. Correction of image deformation from lens distortion, technical report. Technical report, Department of EE and CS, University of Michigan, Ann Arbor, MI, 1987.
- [27] K. D. Gremban, C. E. Thorpe, and T. Kanade. Geometric camera calibration using systems of linear equations. In *IEEE, International Conference on Robotics and Automation*, pages 562–567, 1988.
- [28] J. Hadamard. Sur les problèmes aux érivées partielles et leur signification physique. *Princeton University Bulletin*, 13, 1902.
- [29] G. Healey and R. Kondepudy. CCD camera calibration and noise estimation. In *IEEE, Conference on Computer Vision and Pattern Recognition*, pages 90–95, 1992.
- [30] R.A. Hofmeister. *Amplifying Photoelectric Energy in Low-Light Imaging Devices*, pages 192–194. The Photonics Design and Applications Handbook. Laurin, 1995.
- [31] R. C. Hughes. Accuracy of centroid determination for a high-contrast image with vidicon and solid state array cameras. Technical report, National Research Council Canada, December 1985.
- [32] A. Isaguirre, P. Pu, and J. Summers. A new development in camera calibration: Calibrating a pair of mobile cameras. In *Conference on Robotics and Automation*, pages 74–79, 1985.
- [33] B. Jähne, H. Haußecker, and P. Geißler. *Handbook of Computer Vision and Applications*. Academic Press, 1992.
- [34] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.

- [35] R.K. Lenz and R.Y. Tsai. Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology. *IEEE Transactions on Robotics and Automation*, 7(4):546–551, August 1991.
- [36] S. G. MacLean and H. F. L. Pinkney. Machine vision in space. *Canadian Aeronautics and Space Journal*, 39(2):63–77, June 1993.
- [37] E.J. Martins. Considerations for CCD camera design for machine vision and robotic applications. In *Vision '87, Detroit, Michigan*, June 1987.
- [38] H. A. Martins, J. R. Birk, and R. B. Kelley. Camera models based on data from two calibration planes. *Computer Graphics and Image Processing*, 17:173–180, 1981.
- [39] S. Miller and D. Withey. System architecture for the Space Vision System. In *International Conference on Signal Processing, Applications and Technology, Toronto, Ontario, Canada*, pages 126–130, September 13-15 1998.
- [40] Y. Nomura, M. Sagara, H. Naruse, and A. Ide. Simple calibration algorithm for high-distortion-lens camera. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 14(11):1095–1099, November 1992.
- [41] G. Okouneva. CSVS-04 video characterization: Cohu camera. Technical report, Neptec Design Group Inc., Kanata, Ontario, Canada, April 1999.
- [42] Oriel Instruments Inc. *The book of photon tools*. Publication obtained from www.oriel.com.
- [43] F. Pedersini, A. Sarti, and S. Tubaro. Estimation and compensation of subpixel edge localization error. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(10), November 1997.
- [44] B. Peuchot and M. Saint-André. CCD camera calibration virtual equivalent model. In *IEEE, International Conference on Engineering in Medicine and Biology Society*, pages 1960–1961, 1992.
- [45] P. Puget and T. Skordas. An optimal solution for mobile camera calibration. In *First European Conference on Computer Vision*, pages 187–188, 1990.
- [46] M. Qiu and S. D. Ma. The nonparametric approach for camera calibration. In *IEEE, Fifth International Conference on Computer Vision*, pages 224–229, 1995.

- [47] J. Schneider. *Software Design Document for the Canadian Space Vision System*. Neptec Design Group Inc., Kanata, Ontario, Canada, May 1999.
- [48] S.-W. Shih, Y.-P. Hung, and W.-S. Lin. Accuracy analysis on the estimation of camera parameters for active vision systems. Technical Report TR-IIS-96-006, Institute of Information Science, Academia Sinica, Nankang, Taipei, Taiwan, 1996.
- [49] W.J. Smith. *Optical Components: Optical Objective Systems: A Schematic Guide*, pages 407–411. The Photonics Design and Applications Handbook. Laurin, 1995.
- [50] R.C. Staunton. Edge operator error estimation incorporating measurements of CCD TV camera transfer function. *IEEE Proc.-Vis. Image Signal Processing*, 145(3), June 1998.
- [51] G. P. Stein. Lens distortion calibration using point correspondences. Technical Report A. I. Memo No. 1595, Massachusetts Institute of Technology, November 1996. Publication retrieved by anonymous ftp to publication.ai.mit.edu.
- [52] J.-P. Tarel and A. Gagalowicz. Calibration de caméra à base d’ellipses. Technical Report R.R. no. 2200, Unit de Recherche INRIA Rocquencourt, January 1994.
- [53] J.-P. Tarel and J.-M. Vezien. CamCal v1.0 manual. A complete software solution for camera calibration. Technical Report R.R. no. 0196, Unité de Recherche INRIA Rocquencourt, September 1996.
- [54] Tektronix. *NTSC Video Measurements*, 1997. publication retrieved from: <http://www.tek.com>.
- [55] R.Y. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. In *Conference on Computer Vision and Pattern Recognition*, pages 364–374, 1986.
- [56] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision. *IEEE Transactions on Robotics and Automation*, RA-3(4):323–344, August 1986.
- [57] I. Tsatsakis, E. Kayafas, V. Loumos, and G. Cambourakis. Using low cost video cameras in automation, a close range photogrammetry approach. In *IEEE, International Symposium on Industrial Electronics*, pages 523–528, 1995.

- [58] J.-S. Valois. *Default CLC database v.5*. Neptec Design Group Inc., Kanata, Ontario, Canada, June 1998. camera calibration database for CSVS.
- [59] J.-S. Valois. A new SHC calculation method. Technical report, Neptec Design Group Inc., Kanata, Ontario, Canada, March 2000. Technical report currently under preparation.
- [60] C. Walker. *Primary Lens Aberrations: Seven Ways a Lens Can Blur the Light it's Supposed to Focus*, pages 422–425. The Photonics Design and Applications Handbook. Laurin, 1995.
- [61] D.F. Watson. *Contouring: A guide to the analysis and display of spacial data*. Pergamon, 1994.
- [62] G.-Q. Wei, K. Arbter, and G. Hirzinger. Active self-calibration of robotic eyes and hand-eye relationships with model identification. *IEEE Transactions on Robotics and Automation*, 14(1):158–166, February 1998.
- [63] W.T. Welford. *Aberrations of the Symmetrical Optical System*. Academic Press, 1974.
- [64] D. A. Wells. *Lagrangian Dynamics, Schaum's Outline Series*. McGraw-Hill Book Company, 1967.
- [65] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, October 1992.
- [66] G.A.W West and T.A. Clarke. A survey and examination of subpixel measurement techniques. *SPIE, Close-Range Photogrammetry Meets Machine Vision*, 1(395):465–463, 1990.
- [67] R. G. Willson. *Modeling and Calibration of Automated Zoom Lenses*. PhD thesis, The Robotics Institute at Carnegie Mellon University, January 1994.
- [68] R.G. Willson and S.A. Shafer. What is the center of the image ? Technical report, School of Computer Science, Carnegie Mellon University, April 1993.
- [69] G. Zhou, X. Tang, and B. Yuang. Distortion model selecting and accuracy evaluation of CCD camera calibration. *IEEE Proceedings of ICSP*, pages 875–878, 1996.

APPENDIX A

Camera configurations

The manufacturers specifications for the three cameras and four lenses used for this thesis are presented in this Appendix.

1. CTVC

The Space Shuttle CTVCs are used during the flights to support the crew in performing a number of operations. These include RMS berthing, payload deployment and berthing with and without SVS, EVA support, Cargo bay inspection, Earth observation, and other operations.

Because of their intended use, the camera components are contained in protective, thermally controlled housing. Figure A.1 shows the exposed camera and lens and the flight configuration of a typical CTVC.

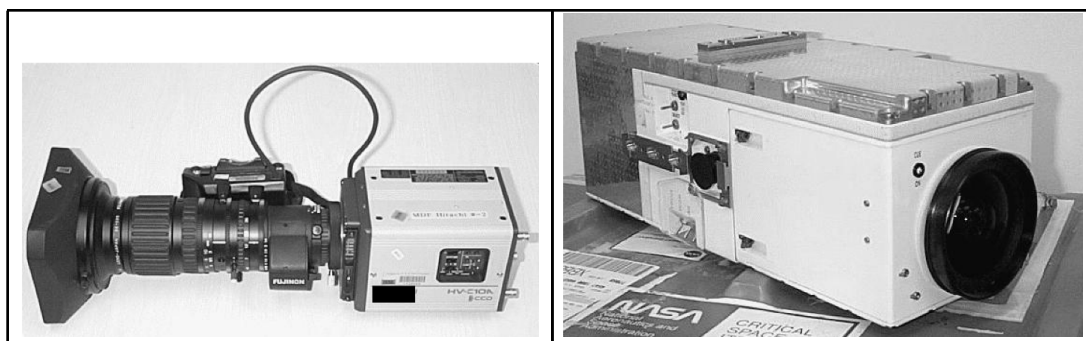


FIGURE A.1. CTVC camera.

Some brief technical specifications on the main components of a CTVC are presented in Tables A.1, and A.2.

Make	All solid-state camera (3 chips), Sanyo Frame Transfer CCD, LC9915G.
Format	2/3", 764H by 487V, 11mm diagonal, Pixel 11.5mm by 13.5mm
Outputs	NTSC composite and component (RGB).
Resolution	100% MTF at TVL/PH

TABLE A.1. CTVC imager properties.


Make	Fujinon F/1.7 5.5-47mm lens, 4:3 aspect ratio.
Zoom	8.5:1, 13.3° to 90°, 840 steps, (90° rotation, 0.11°/step).
Focus	Infinity to 0.3 meter (11.8 in.), (1200 steps, 144° rotation, 0.12°/step).
Iris	F/1.7 to F/16 to Iris fully closed, (360 steps, 76° rotation, 0.21°/step).
Lens diagram	

TABLE A.2. CTVC lens properties.

The horizontal field of view (HFOV) used for the calibrations was 36.4°, which is a typical zoom setting for a SVS configuration. The details on other camera/lens adjustments used for the tests are presented in Table A.3.

Make	Martin Marietta, closed circuit solid state Color Television Camera (CTVC).
s/n	217
Gain	12 dB
Balance	CABIN
Focus	333 feet
Gamma	1 (linear)
Iris	1T (fully open)
Zoom	36.4°
Focal length	13.5 (VITS DATA)

TABLE A.3. CTVC camera configuration.

2. ITVC

The Space Shuttle ITVCs were introduced by the Space Shuttle program to perform operations in dark lighting conditions. The ITVCs are equipped with the so-called ICCDs (Intensified Charged Coupled Devices) that amplify the amount of light perceived.

The term “image intensifier” refers to a series of special imaging tubes that have been designed to amplify, by several order of magnitudes, the emissions of photons falling upon a photocathode. The technology allows the imaging device to obtain more information about the scene being viewed than would be possible under the same lighting conditions without the amplification. Valuable information on the ICCD technology can be found in [30].

In its final configuration the ITVC looks similar to the CTVC. Figure A.2 shows the Angénieux lens used with the ITVCs as well as an ITVC in its flight configuration.

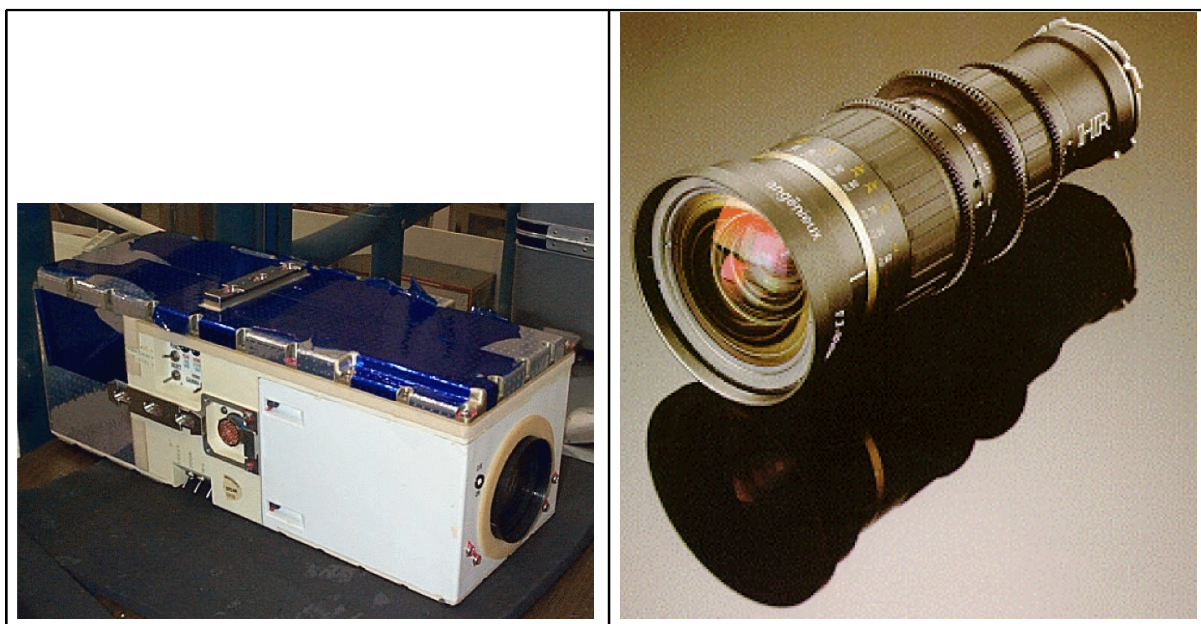


FIGURE A.2. ITVC camera.

Some brief technical specifications on the main components of a ITVC are presented in Tables A.4 and A.5.

Make	Sanyo frame transfer CCD, LC9915G-NS04
Format	2/3", 768H by 490V, 11mm diagonal, Pixel 11.5mm by 13.5mm
Output	EIA RS-170 Composite 75 Ω balanced differential.
Resolution	14% MTF at 345 TVL/PH w/aperture comp., lens T/4.7
Intensifier	ITT Proximity-focused diode intensifier (PFD), 18mm input diagonal, P20 phosphor.

TABLE A.4. ITVC imager properties.

Make	Angénieux, 15 x 8.3 AIF.HR, 4:3 aspect ratio.
Zoom	15:1, 8.3 - 125 mm, 56° - 4°.
Focus	Infinity to 0.8 m (31").
Iris	F/1.7 to fully closed.

TABLE A.5. ITVC lens properties.

The horizontal field of view (HFOV) used for the calibrations was 39.1°, which is a typical zoom setting for a SVS configuration. The details on other camera/lens adjustments used for the tests are presented in Table A.6.

Make	Martin Marietta, closed circuit solid state Intensified Television Camera (ITVC).
s/n	814
Gain	12 dB
Balance	CABIN
Focus	227.5 feet
Gamma	1 (linear)
Iris	12T (fully open)
Zoom	39.1°
Focal length	12.4 (VITS DATA)

TABLE A.6. ITVC camera configuration.

3. Cohu 4910

Cohu is a company that offers a wide variety of high-quality color and monochrome cameras for various types of applications. For the purpose of our tests the model 4910 (shown in Figure A.3) was selected together with two different Pentax/Cosmicar C-mount lenses. A generic representation of the lenses is also present in Figure A.3.

Tables A.7, A.8, and A.9 contain technical information on the Cohu camera and the two lenses used for the tests.

Make	Cohu 4910, High Performance Monochrome Interline Transfer CCD Camera.
Format	1/2", 6.4 by 4.8 mm, 768 (H) by 494 (V)
Output	NTSC composite, unbalanced.
Resolution	580 horizontal, 350 vertical

TABLE A.7. Cohu camera imager properties.

The details on the setting of the Cohu camera during the tests are presented in Table A.10.

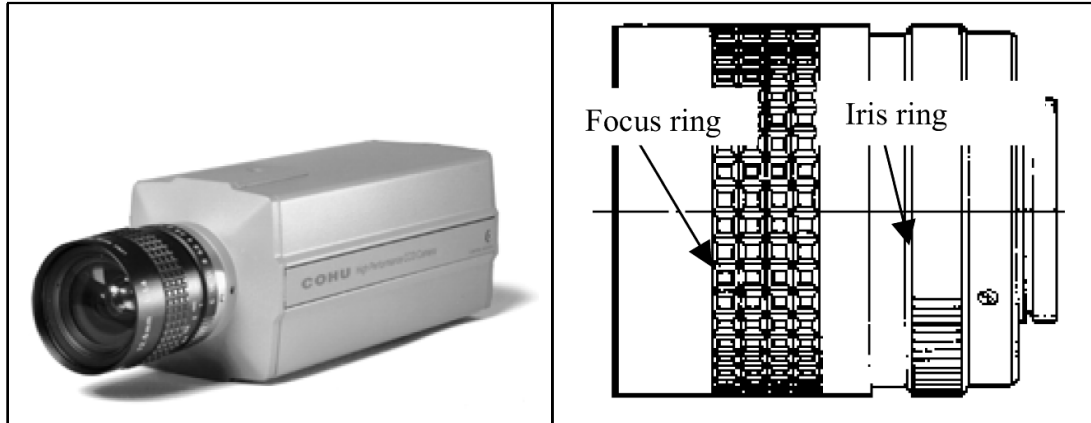


FIGURE A.3. Cohu camera and Cosmocar lens.

Make	Cosmicar/Pentax C60606.
Type	Fixed focal length.
Focus	Infinity to 0.3 meters.
Iris	8T to fully open.

TABLE A.8. Cosmicar 6 mm lens properties.

Make	Cosmicar/Pentax C30808.
Type	Fixed focal length.
Focus	Infinity to 0.3 meters.
Iris	8T to fully open.

TABLE A.9. Cosmicar 8 mm lens properties.

Make	Cohu 4910
s/n	235038
Gain	0 dB
Balance	Full CCW (hard stop)
Shutter	Off
Peak AVG.	Default
AGC	Off
Gamma	1 (linear)
Sharpness	Default

TABLE A.10. Cohu camera configuration.

APPENDIX B

Least-squares minimization method

This appendix contains the details of the Least-Squares minimization techniques used to solve the non-linear parametric calibration equations. The method is explained through the *Radial - first order* implementation found in Chapter 5, Section 2.3.1.

1. Radial - first order example

The colinear relations of photogrammetry from Section 2 are used to derive the mathematical expressions to minimize χ by the LS fitting criterion for a total number of N target :

$$\chi = \arg \min_{\vec{x}} \left[\sum_{j=1}^N (\Delta_y^2(j) + \Delta_z^2(j)) \right],$$

where:

$$\begin{aligned} \Delta_y(j) &= y_U(j) (X_{CT} + X_T \alpha_{11} + Y_T \alpha_{12} + Z_T \alpha_{13}) \\ &\quad - f_e (Y_{CT} + X_T \alpha_{21} + Y_T \alpha_{22} + Z_T \alpha_{23}), \\ \Delta_z(j) &= z_U(j) (X_{CT} + X_T \alpha_{11} + Y_T \alpha_{12} + Z_T \alpha_{13}) \\ &\quad - f_e (Z_{CT} + X_T \alpha_{31} + Y_T \alpha_{32} + Z_T \alpha_{33}). \end{aligned} \tag{B.1}$$

The previous equation can be simplified to:

$$\begin{aligned}
X'_T &= X_T \alpha_{11} + Y_T \alpha_{12} + Z_T \alpha_{13}, \\
Y'_T &= X_T \alpha_{21} + Y_T \alpha_{22} + Z_T \alpha_{23}, \\
Z'_T &= X_T \alpha_{31} + Y_T \alpha_{32} + Z_T \alpha_{33}.
\end{aligned}$$

The corrected centroids (y_U, z_U) in the ICS are obtained from the distorted centroids (y_F, z_F) in the FCS, using the following transformations:

$$\begin{aligned}
y_U &= y_d \cdot (k_1 \cdot (y_d^2 + z_d^2) + 1), \\
z_U &= z_d \cdot (k_1 \cdot (y_d^2 + z_d^2) + 1),
\end{aligned}$$

where

$$\begin{aligned}
y_d &= (y_F - C_y) \cdot G_y, \\
z_d &= (z_F - C_z) \cdot G_z.
\end{aligned}$$

The unknown parameters in the previous equations are: G_y , G_z , C_y , C_z , K_1 , X_{CT} , Y_{CT} , Z_{CT} , α_{12} , α_{13} , and α_{23} .

A first order Taylor's series expansion is used to linearized Equation B.2. The coefficients are obtained by applying the following equations:

$$\begin{aligned}
B_{11} : \frac{\partial \Delta_y(j)}{\partial G_y} &= (X_{CT} + X'_T) \cdot \left[\frac{y_d + 3k_1 y_d^3 + k_1 y_d z_d^2}{G_y} \right], \\
B_{12} : \frac{\partial \Delta_y(j)}{\partial G_z} &= (X_{CT} + X'_T) \cdot \left[\frac{2k_1 y_d z_d^2}{G_z} \right], \\
B_{13} : \frac{\partial \Delta_y(j)}{\partial C_y} &= (X_{CT} + X'_T) \cdot [-G_y - 3G_y k_1 y_d^2 - G_y k_1 z_d^2], \\
B_{14} : \frac{\partial \Delta_y(j)}{\partial C_z} &= (X_{CT} + X'_T) \cdot [-2G_z k_1 y_d z_d], \\
B_{15} : \frac{\partial \Delta_y(j)}{\partial k_1} &= (X_{CT} + X'_T) \cdot [y_d^3 + y_d z_d^2],
\end{aligned}$$

$$B_{16} : \frac{\partial \Delta_y(j)}{\partial X_{CT}} = y_u(j),$$

$$B_{17} : \frac{\partial \Delta_y(j)}{\partial Y_{CT}} = -f_e,$$

$$B_{18} : \frac{\partial \Delta_y(j)}{\partial Z_{CT}} = 0,$$

$$B_{19} : \frac{\partial \Delta_y(j)}{\partial \alpha_{12}} = y_u(j) \cdot \left[\frac{\alpha_{33} Y'_T(j)}{\alpha_{11}} - \frac{\alpha_{23} Z'_T(j)}{\alpha_{11}} \right] - f_e \cdot \left[\frac{\alpha_{13} Z'_T(j)}{\alpha_{11}} - \frac{\alpha_{33} X'_T(j)}{\alpha_{11}} \right],$$

$$B_{110} : \frac{\partial \Delta_y(j)}{\partial \alpha_{13}} = y_u(j) \cdot \left[\frac{-\alpha_{32} Y'_T(j)}{\alpha_{11}} + \frac{\alpha_{22} Z'_T(j)}{\alpha_{11}} \right] - f_e \cdot \left[\frac{\alpha_{32} X'_T(j)}{\alpha_{11}} - \frac{\alpha_{32} \alpha_{13} Z'_T(j)}{\alpha_{11} \alpha_{33}} \right],$$

$$B_{111} : \frac{\partial \Delta_y(j)}{\partial \alpha_{23}} = -f_e \cdot \left[\frac{Z'_T(j)}{\alpha_{33}} \right],$$

$$B_{21} : \frac{\partial \Delta_z(j)}{\partial G_y} = (X_{CT} + X'_T) \cdot \left[\frac{2k_1 z_d y_d^2}{G_y} \right],$$

$$B_{22} : \frac{\partial \Delta_z(j)}{\partial G_z} = (X_{CT} + X'_T) \cdot \left[\frac{z_d + 3k_1 z_d^3 + k_1 y_d^2 z_d}{G_y} \right],$$

$$B_{23} : \frac{\partial \Delta_z(j)}{\partial C_y} = (X_{CT} + X'_T) \cdot [-2k_1 y_d z_d],$$

$$B_{24} : \frac{\partial \Delta_z(j)}{\partial C_z} = (X_{CT} + X'_T) \cdot [-G_y - 3G_y k_1 z_d^2 - G_z k_1 y_d^2],$$

$$B_{25} : \frac{\partial \Delta_z(j)}{\partial k_1} = (X_{CT} + X'_T) \cdot [y_d^2 z_d + z_d^3],$$

$$B_{26} : \frac{\partial \Delta_z(j)}{\partial X_{CT}} = z_u(j),$$

$$B_{27} : \frac{\partial \Delta_z(j)}{\partial Y_{CT}} = 0,$$

$$B_{28} : \frac{\partial \Delta_z(j)}{\partial Z_{CT}} = -f_e,$$

$$B_{29} : \frac{\partial \Delta_z(j)}{\partial \alpha_{12}} = z_u(j) \cdot \left[\frac{\alpha_{33} Y'_T(j)}{\alpha_{11}} - \frac{\alpha_{23} Z'_T(j)}{\alpha_{11}} \right] - f_e \cdot \left[\frac{\alpha_{23} X'_T(j)}{\alpha_{11}} - \frac{\alpha_{13} Y'_T(j)}{\alpha_{11}} \right],$$

$$B_{210} : \frac{\partial \Delta_z(j)}{\partial \alpha_{13}} = z_u(j) \cdot \left[\frac{-\alpha_{32} Y'_T(j)}{\alpha_{11}} + \frac{\alpha_{22} Z'_T(j)}{\alpha_{11}} \right] - f_e \cdot \left[\frac{-\alpha_{22} X'_T(j)}{\alpha_{11}} - \frac{\alpha_{32} \alpha_{13} Y'_T(j)}{\alpha_{11} \alpha_{33}} \right],$$

$$B_{211} : \frac{\partial \Delta_z(j)}{\partial \alpha_{23}} = -f_e \cdot \left[\frac{X'_T(j)}{\alpha_{33}} \right].$$

Rewriting the linearized version of Equation B.2, yields:

$$\begin{bmatrix} \Delta_y(j) \\ \Delta_z(j) \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} & B_{15} & B_{16} & B_{17} & B_{18} & B_{18} & B_{110} & B_{111} \\ B_{21} & B_{22} & B_{23} & B_{24} & B_{25} & B_{26} & B_{27} & B_{28} & B_{28} & B_{210} & B_{211} \end{bmatrix} \cdot [\Delta G_y \Delta G_z \Delta C_y \Delta C_z \Delta K_1 \Delta X_{CT} \Delta Y_{CT} \Delta Z_{CT} \Delta \alpha_{12} \Delta \alpha_{13} \Delta \alpha_{23}]^T = 0.$$

The value for the Δ vector is obtained by the pseudo-inverse method, as follows:

$$\Delta = (B^T \cdot B)^{-1} \cdot B^T \cdot (-U).$$

Figure B.1 contains a flow chart describing the iterative process followed by this minimization method.

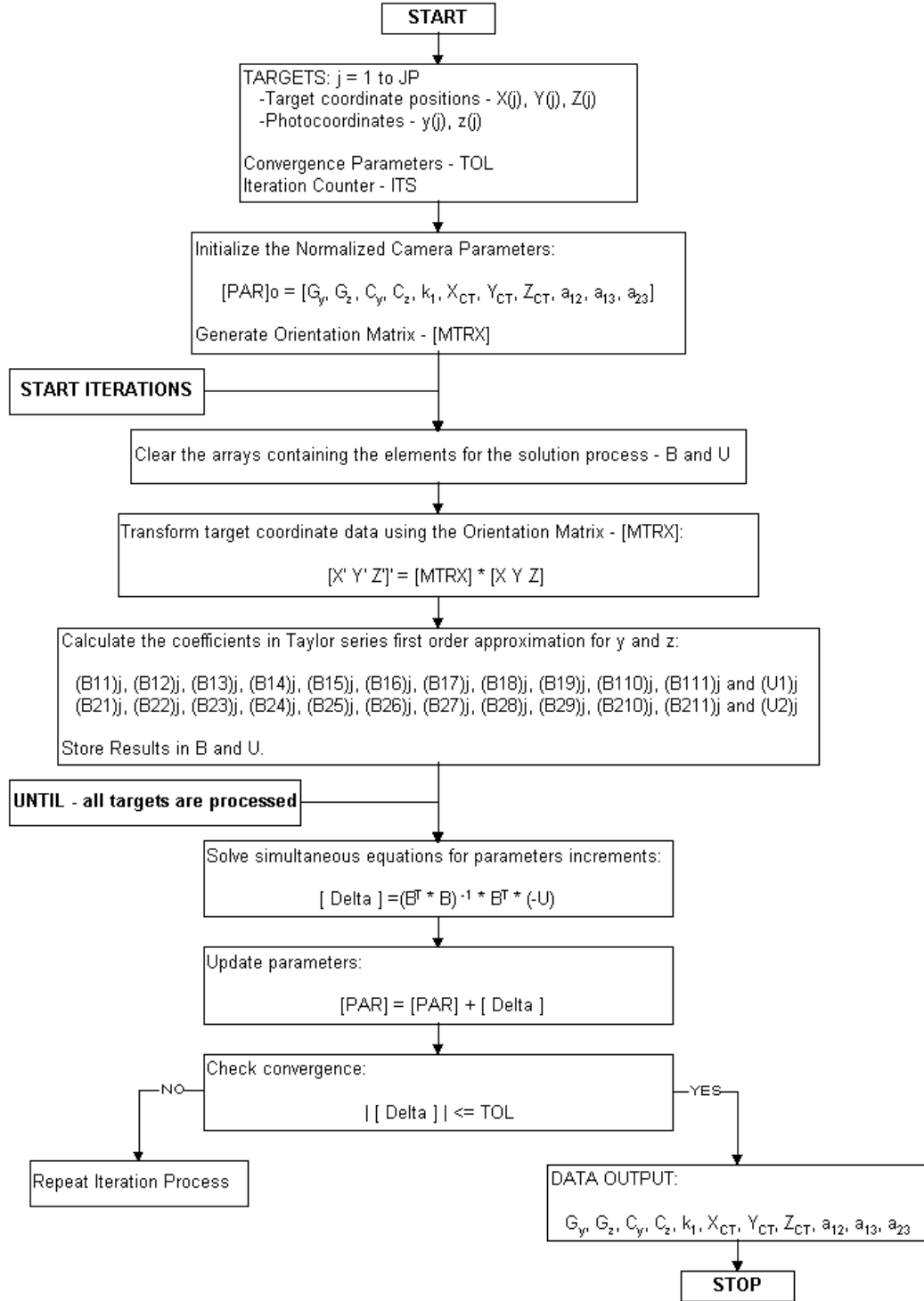


FIGURE B.1. Least-Squares minimization flow chart.

APPENDIX C

Student's t-test results

The complete results of the Student's t-tests from Chapter 7 are presented in this appendix. The data is compiled for every test and camera in various tables which include all methods, group numbers, and average errors.

1. Student's t-test in a nutshell

For two normally distributed data sets of n_1 and n_2 elements and given the means ($\overline{X_1}$ and $\overline{X_2}$) and variance (σ_1^2 and σ_2^2) of both sets, the *t-value* is obtained by applying Formula C.1.

$$(C.1) \quad t - value = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{\sigma_1^2}{n_1-1} + \frac{\sigma_2^2}{n_2-1}}}.$$

The level of significance is determined by the *t-value* and the number of elements in both sets. Significance levels of 95% and above indicate that the mean difference between the two distributions are important enough to consider the two sets different.

2. Complete calibration performance results

This section presents the calibration performance results obtained for the different cameras and tests.

Groups	Average error (lines)	Methods
1	0.0437	◊ Lookup compensation method
	0.0448	◊ Polynomial smoothing method with Lookup compensation method
2	0.0499	◊ Radial - fixed gain ratios
	0.0534	◊ Radial and decentering - 2 steps
	0.0543	◊ Radial - fixed focal length
3	0.0578	◊ Radial and decentering
	0.0589	◊ Averaging of multiple maps
	0.0604	◊ Original method
	0.0626	◊ Polynomial smoothing method
	0.0627	◊ Special radial
	0.0633	◊ Radial and second order
	0.0648	◊ Radial and thin prism
4	0.0681	◊ Radial - different data sets
	0.0685	◊ Radial and tangential
	0.0734	Standard Centroid Error

TABLE C.1. Calibration performance from the *CSE* on the calibration board for the CTVC.

Groups	Average error (lines)	Methods
1	0.0720	◊ Lookup compensation method
	0.0736	Standard Centroid Error
	0.0754	◊ Original method
2	0.0889	◊ Averaging of multiple maps
3	0.1043	◊ Polynomial smoothing method with Lookup compensation method
	0.1077	◊ Polynomial smoothing method
4	0.4668	◊ Radial - fixed gain ratios
	0.4939	◊ Radial - fixed focal length
5	0.5668	◊ Radial - different data sets
	0.5714	◊ Radial and second order
	0.5775	◊ Radial and decentering
	0.5779	◊ Special radial
	0.5804	◊ Radial and thin prism
6	DNC	◊ Radial and decentering - 2 steps
	DNC	◊ Radial and tangential

TABLE C.2. Calibration performance from the *CSE* on the calibration board for the ITVC. (DNC: calibration did not converge)

Groups	Average error (lines)	Methods
1	0.0537	◇ Lookup compensation method
	0.0583	◇ Polynomial smoothing method with Lookup compensation method
	0.0733	Standard Centroid Error
2	0.0906	◇ Averaging of multiple maps
3	0.1196	◇ Polynomial smoothing method
	0.1220	◇ Original method
4	0.1783	◇ Radial and second order
	0.1787	◇ Radial and decentering
	0.1789	◇ Radial and decentering - 2 steps
	0.1837	◇ Radial and tangential
	0.1889	◇ Radial and thin prism
	0.1901	◇ Special radial
	0.1914	◇ Radial - fixed gain ratios
	0.1918	◇ Radial - fixed focal length
	0.2003	◇ Radial - different data sets

TABLE C.3. Calibration performance from the *CSE* on the calibration board for the Cohu with a 6 mm lens.

Groups	Average error (lines)	Methods
1	0.0478	◇ Polynomial smoothing method with Lookup compensation method
	0.0482	◇ Averaging of multiple maps
	0.0489	◇ Lookup compensation method
2	0.0619	◇ Polynomial smoothing method
	0.0624	◇ Original method
	0.0733	Standard Centroid Error
3	0.0795	◇ Radial and second order
	0.0842	◇ Radial and decentering - 2 steps
	0.0843	◇ Radial and decentering
	0.0849	◇ Radial and thin prism
	0.0868	◇ Radial - different data sets
	0.0869	◇ Radial - fixed focal length
	0.0870	◇ Radial - fixed gain ratios
	0.0870	◇ Special radial
	0.0894	◇ Radial and tangential

TABLE C.4. Calibration performance from the *CSE* on the calibration board for the Cohu with an 8 mm lens.

Groups	Average error (lines)	Methods
	0.0624	Standard Centroid Error
1	0.2859 0.3970	◇ Radial and decentering ◇ Radial and tangential
2	0.4082 0.4308 0.4314 0.4456 0.4465 0.4476 0.4542 0.4545 0.4546 0.4553 0.4562 0.4645	◇ Radial - different data sets ◇ Polynomial smoothing method with Lookup compensation method ◇ Lookup compensation method ◇ Original method ◇ Polynomial smoothing method ◇ Radial and decentering - 2 steps ◇ Radial and thin prism ◇ Radial - fixed focal length ◇ Special radial ◇ Radial and second order ◇ Radial - fixed gain ratios ◇ Averaging of multiple maps

TABLE C.5. Calibration performance from the *CSE* on the volumetric array for the CTVC.

Groups	Average error (lines)	Methods
	0.0618	Standard Centroid Error
1	0.3879 0.3908	◇ Original method ◇ Lookup compensation method
2	0.4007 0.4078 0.4089 0.6386 0.6552 0.6803 0.6848 0.6983 0.7026 0.7138	◇ Averaging of multiple maps ◇ Polynomial smoothing method ◇ Polynomial smoothing method with Lookup compensation method ◇ Radial and second order ◇ Special radial ◇ Radial and thin prism ◇ Radial and decentering ◇ Radial - fixed focal length ◇ Radial - fixed gain ratios ◇ Radial - different data sets
3	DNC DNC	◇ Radial and decentering - 2 steps ◇ Radial and tangential

TABLE C.6. Calibration performance from the *CSE* on the volumetric array for the ITVC. (DNC: calibration did not converge)

Groups	Average error (lines)	Methods
	0.0676	Standard Centroid Error
1	0.252	◇ Radial and tangential
	0.2544	◇ Polynomial smoothing method with Lookup compensation method
	0.2573	◇ Special radial
	0.2589	◇ Polynomial smoothing method
	0.2594	◇ Radial - fixed gain ratios
	0.2611	◇ Original method
	0.2630	◇ Lookup compensation method
	0.2636	◇ Averaging of multiple maps
	0.2707	◇ Radial - different data sets
	0.2710	◇ Radial and second order
	0.2717	◇ Radial and decentering
	0.2718	◇ Radial and decentering - 2 steps
	0.2722	◇ Radial and thin prism
2	0.3385	◇ Radial - fixed focal length

TABLE C.7. Calibration performance from the *CSE* on the volumetric array for the Cohu with a 6 mm lens.

Groups	Average error (lines)	Methods
	0.0638	Standard Centroid Error
1	0.2435	◇ Original method
	0.2471	◇ Polynomial smoothing method
	0.2521	◇ Lookup compensation method
	0.2572	◇ Polynomial smoothing method with Lookup compensation method
	0.2589	◇ Radial and tangential
	0.2600	◇ Averaging of multiple maps
	0.2660	◇ Radial and decentering - 2 steps
	0.2667	◇ Radial and thin prism
	0.2681	◇ Radial and second order
	0.2697	◇ Radial and decentering
	0.2699	◇ Radial - different data sets
	0.2720	◇ Radial - fixed focal length
	0.2735	◇ Special radial
	0.2742	◇ Radial - fixed gain ratios

TABLE C.8. Calibration performance from the *CSE* on the volumetric array for the Cohu with an 8 mm lens.

Groups	Average error (lines)	Methods
	0.0197	Standard Centroid Error
1	0.1105	◇ Radial and tangential
	0.1130	◇ Radial - different data sets
	0.1170	◇ Lookup compensation method
	0.1172	◇ Polynomial smoothing method with Lookup compensation method
	0.1215	◇ Original method
	0.1220	◇ Polynomial smoothing method
	0.1236	◇ Radial and decentering - 2 steps
	0.1240	◇ Radial and decentering
	0.1252	◇ Radial and thin prism
	0.1252	◇ Special radial
	0.1255	◇ Radial and second order
	0.1256	◇ Radial - fixed focal length
	0.1257	◇ Radial - fixed gain ratios
	0.1308	◇ Averaging of multiple maps

TABLE C.9. Calibration performance from the *OSE* on the volumetric array for the CTVC.

Groups	Average error (lines)	Methods
	0.0192	Standard Centroid Error
1	0.1035	◇ Original method
	0.1040	◇ Lookup compensation method
	0.1071	◇ Averaging of multiple maps
	0.1083	◇ Polynomial smoothing method
	0.1085	◇ Polynomial smoothing method with Lookup compensation method
2	0.1751	◇ Radial and second order
	0.1805	◇ Special radial
	0.1877	◇ Radial and thin prism
	0.1882	◇ Radial - fixed focal length
	0.1886	◇ Radial and decentering
	0.1926	◇ Radial - fixed gain ratios
	0.1957	◇ Radial - different data sets
3	DNC	◇ Radial and decentering - 2 steps
	DNC	◇ Radial and tangential

TABLE C.10. Calibration performance from the *OSE* on the volumetric array for the ITVC. (DNC: calibration did not converge)

Groups	Average error (lines)	Methods
	0.0247	Standard Centroid Error
1	0.0806	◇ Radial and tangential
	0.0818	◇ Special radial
	0.0828	◇ Radial - fixed gain ratios
	0.0843	◇ Polynomial smoothing method with Lookup compensation method
	0.0857	◇ Polynomial smoothing method
	0.0863	◇ Original method
	0.0868	◇ Radial and thin prism
	0.0868	◇ Radial and second order
	0.0869	◇ Averaging of multiple maps
	0.0870	◇ Radial and decentering
	0.0870	◇ Radial and decentering - 2 steps
	0.0871	◇ Lookup compensation method
	0.0878	◇ Radial - different data sets
2	0.1119	◇ Radial - fixed focal length

TABLE C.11. Calibration performance from the *OSE* on the volumetric array for the Cohu with a 6 mm lens.

Groups	Average error (lines)	Methods
	0.0207	Standard Centroid Error
1	0.0711	◇ Original method
	0.0719	◇ Polynomial smoothing method
	0.0734	◇ Lookup compensation method
	0.0745	◇ Polynomial smoothing method with Lookup compensation method
	0.0753	◇ Averaging of multiple maps
	0.0756	◇ Radial and tangential
	0.0771	◇ Radial and decentering - 2 steps
	0.0772	◇ Radial and thin prism
	0.0776	◇ Radial and second order
	0.0778	◇ Radial - different data sets
	0.0780	◇ Radial and decentering
	0.0789	◇ Radial - fixed focal length
	0.0793	◇ Special radial
	0.0795	◇ Radial - fixed gain ratios

TABLE C.12. Calibration performance from the *OSE* on the volumetric array for the Cohu with an 8 mm lens.

Groups	Average error	Methods
Translation error (inches)		
1	0.7977	◇ Averaging of multiple maps
	0.8051	◇ Radial - different data sets
	0.8078	◇ Radial and tangential
	0.8080	◇ Radial - fixed focal length
	0.8087	◇ Radial and decentering - 2 steps
	0.8133	◇ Radial - fixed gain ratios
	0.8159	◇ Radial and decentering
	0.8194	◇ Original method
	0.8279	◇ Polynomial smoothing method
	0.8296	◇ Lookup compensation method
	0.8342	◇ Radial and thin prism
	0.8356	◇ Polynomial smoothing method with Lookup compensation method
	0.8369	◇ Radial and second order
	0.8401	◇ Special radial
Rotation error (degrees)		
	0.1329	◇ Radial and tangential
	0.1367	◇ Radial - different data sets
	0.1391	◇ Lookup compensation method
	0.1402	◇ Polynomial smoothing method with Lookup compensation method
	0.1500	◇ Original method
	0.1517	◇ Polynomial smoothing method
	0.1553	◇ Radial and decentering - 2 steps
	0.1561	◇ Radial and decentering
	0.1569	◇ Radial and second order
	0.1576	◇ Special radial
	0.1589	◇ Radial and thin prism
	0.1589	◇ Radial - fixed focal length
	0.1597	◇ Radial - fixed gain ratios
	0.1755	◇ Averaging of multiple maps

TABLE C.13. Calibration performance from the *RPSE* on the volumetric array for the CTVC.

Groups	Average error	Methods
Translation error (inches)		
1	0.4074 0.6770	◊ Radial - fixed focal length ◊ Special radial
2	0.7941 0.7963 0.8046 0.8101 0.8197 0.8442 0.8701 0.9742 1.0290 1.0501	◊ Radial - different data sets ◊ Polynomial smoothing method ◊ Radial and second order ◊ Polynomial smoothing method with Lookup compensation method ◊ Radial and thin prism ◊ Radial and decentering ◊ Radial - fixed gain ratios ◊ Averaging of multiple maps ◊ Original method ◊ Lookup compensation method
3	DNC ¹ DNC	◊ Radial and decentering - 2 steps ◊ Radial and tangential
Rotation error (degrees)		
1	0.0902 0.0920 0.0951 0.0991 0.1039 0.1070	◊ Radial - fixed focal length ◊ Polynomial smoothing method with Lookup compensation method ◊ Polynomial smoothing method ◊ Lookup compensation method ◊ Original method ◊ Averaging of multiple maps
2	0.2382 0.2490 0.2591 0.2617 0.2652 0.2652	◊ Radial and second order ◊ Radial and decentering ◊ Special radial ◊ Radial and thin prism ◊ Radial - different data sets ◊ Radial - fixed gain ratios
3	DNC DNC	◊ Radial and decentering - 2 steps ◊ Radial and tangential

TABLE C.14. Calibration performance from the *RPSE* on the volumetric array for the ITVC. (DNC: calibration did not converge)

Groups	Average error	Methods
Translation error (inches)		
	0.3543	◇ Radial and tangential
	0.3601	◇ Radial - different data sets
	0.3689	◇ Radial - fixed gain ratios
	0.3758	◇ Radial - fixed focal length
	0.3776	◇ Special radial
	0.3874	◇ Radial and decentering
	0.3901	◇ Radial and second order
	0.3901	◇ Radial and decentering - 2 steps
	0.3914	◇ Radial and thin prism
	0.4444	◇ Polynomial smoothing method with Lookup compensation method
	0.4461	◇ Original method
	0.4486	◇ Lookup compensation method
	0.4500	◇ Polynomial smoothing method
	0.4677	◇ Averaging of multiple maps
Rotation error (degrees)		
	0.1119	◇ Special radial
	0.1139	◇ Radial - fixed gain ratios
	0.1154	◇ Radial - fixed focal length
	0.1178	◇ Radial and tangential
	0.1186	◇ Radial and second order
	0.1201	◇ Radial and decentering - 2 steps
	0.1204	◇ Radial and decentering
	0.1208	◇ Radial and thin prism
	0.1217	◇ Radial - different data sets
	0.1258	◇ Polynomial smoothing method
	0.1261	◇ Polynomial smoothing method with Lookup compensation method
	0.1291	◇ Averaging of multiple maps
	0.1323	◇ Lookup compensation method
	0.1332	◇ Original method

TABLE C.15. Calibration performance from the *RPSE* on the volumetric array for the Cohu with a 6 mm lens.

Groups	Average error	Methods
Translation error (inches)		
	0.3242	◇ Original method
	0.3306	◇ Polynomial smoothing method
	0.3533	◇ Lookup compensation method
	0.3607	◇ Polynomial smoothing method with Lookup compensation method
	0.3771	◇ Radial - different data sets
	0.3780	◇ Radial and decentering - 2 steps
	0.3790	◇ Radial and thin prism
	0.3841	◇ Radial and decentering
	0.3853	◇ Radial - fixed focal length
	0.3866	◇ Radial and tangential
	0.3866	◇ Averaging of multiple maps
	0.3886	◇ Special radial
	0.3887	◇ Radial - fixed gain ratios
	0.3909	◇ Radial and second order
Rotation error (degrees)		
	0.0938	◇ Original method
	0.0978	◇ Polynomial smoothing method
	0.1000	◇ Lookup compensation method
	0.1046	◇ Polynomial smoothing method with Lookup compensation method
	0.1069	◇ Averaging of multiple maps
	0.1162	◇ Radial and second order
	0.1185	◇ Radial - different data sets
	0.1197	◇ Radial and decentering - 2 steps
	0.1203	◇ Radial and thin prism
	0.1205	◇ Radial and tangential
	0.1210	◇ Radial and decentering
	0.1238	◇ Radial - fixed focal length
	0.1243	◇ Radial - fixed gain ratios
	0.1244	◇ Special radial

TABLE C.16. Calibration performance from the *RPSE* on the volumetric array for the Cohu with an 8 mm lens.

Document Log:

Manuscript Version 1

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$ - $\text{\texttt{L}}\text{\texttt{A}}\text{\texttt{T}}\text{\texttt{E}}\text{\texttt{X}}$ — 21 June 2000

JEAN-SÉBASTIEN VALOIS

CENTER FOR INTELLIGENT MACHINES, MCGILL UNIVERSITY, 3480 UNIVERSITY ST., MONTRÉAL
(QUÉBEC) H3A 2A7, CANADA, *Tel.* : (514) 398-2185

E-mail address: `valois@cim.mcgill.ca`

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$ - $\text{\texttt{L}}\text{\texttt{A}}\text{\texttt{T}}\text{\texttt{E}}\text{\texttt{X}}$