# THE QUESTION OF ACCURACY IN CAMERA CALIBRATION AND A NOVEL TECHNIQUE FOR ESTIMATING CAMERA GEOMETRY

SWAPNA PUTHUKKUDICHALIL

# The Question of Accuracy in Camera Calibration and a Novel Technique for Estimating Camera Geometry

© Swapna Puthukkudichalil, B. Tech.

A thesis submitted to the
School of Graduate Studies in partial
fulfilment of the requirements for the degree of
Master of Engineering

Faculty of Engineering and Applied Science
Memorial University of Newfoundland

August 2009

St. John's Newfoundland

**Abstract**

Machine Vision is one of the leading technologies of the present era. The quest for the best solution to the geometrical camera calibration has been a research problem for many years in the Computer/Machine Vision communities. In the field of machine vision, camera calibration refers to the experimental determination of a set of parameters which describe the image formation process for a given analytical model of the machine vision system. A complete set of calibration parameters includes both the intrinsic parameters that describe the lens-camera-frame grabber combination as well as the extrinsic parameters that relate the position and orientation of the camera to a fixed reference frame. An accurate, reliable calibration procedure is essential for most industrial machine vision applications including mechanical metrology, robot assembly, reverse engineering, tracking, image fusion etc. One of the most systematic calibration procedures for 3-D machine vision applications was proposed by Heikkila in which a comprehensive set of camera parameters is automatically evaluated by observing a calibration target consisting of two perpendicular planes, each with 256 circular control points. Other similar techniques employ a checkerboard pattern as a target and use the vertices of the squares as control points. While these techniques are sound from a theoretical point of view, they do not adequately speak to the question of measurement accuracy. This study addresses the problems associated with geometric camera calibration using the Design of Experiments as well as the question of accuracy. The broad use of camera calibration in the field of machine vision and computer vision is the primary motivation for understanding the major factors that influence the accuracy of camera calibration as well as the accuracy that can be achieved within the practical limits of most research environments. Based on the results of this study, a new approach towards camera calibration is also proposed in which image center and focal length are calculated independently of the lens distortion.

**Acknowledgements**

The author owes a debt of sincere gratitude and thanks to all who helped and encouraged her to make this research work a success. Firstly, the author would like to express her sincere and profound sense of gratitude and respect for her research supervisors Dr. Nicholas Krouglicof (main supervisor) and Dr. Ray Gosine (co-supervisor), Faculty of Engineering and Applied Science, Memorial University of Newfoundland (MUN), St. John's, Canada. The author gratefully acknowledges Dr. Nicholas Krouglicof for his valuable guidance, advice and dedicated help given at every stage of the research work. The author indeed expresses her sincere thanks for his constant encouragement, untiring support and constructive suggestions, which helped in improving the quality of this research work.

The author also wishes to express her sincere thanks and respect for all her other teachers in MUN who taught and educated her during courses.

The author is extremely grateful to the Intelligent Systems Lab for providing the facilities and work environment necessary to complete this research work.

The author gratefully acknowledges her colleague Mr. Taufiqur Rahman, for his co-operation, help and support shown during certain stages of the research work. The author is also grateful to her colleague Mr. Migara Liyanage and to all other fellow scholars/friends in the Intelligent Systems Lab for their help and support during her program of study.

The author wishes to express sincere gratitude for her father Mr. P. Subramanian, mother Mrs. Valli A., sister Ms. Dilna P. for their prayers, moral support, encouragement, unconditional love and affection. The author also gratefully remembers and respects her beloved grandparents for their love and affection. Finally, the author would like to express her sincere thanks to all her family members, near and dear ones, close relatives and friends who supported her in no small measure to accomplish this target.

Last, but certainly not least, the author is extremely grateful and thankful to her husband Mr. Premkumar for his prayers, patience, help, immeasurable love and affection, moral support, constant encouragement and inspiration shown throughout her research life.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations and Symbols

| | |
|---|---|
| $\phi$ | Direction of the sobel gradient image |
| $F(a_d, \delta)$ | Correction for the combined effect of radial and tangential distortion |
| $\theta$ | Angle |
| $(\overline{X}, \overline{Y})$ | Centroid of the image |
| $\varepsilon$ | Error estimate |
| $\theta_{ext}$ | Extrinsic parameters |
| $\sigma$ | Gaussian filter |
| $\prod$ | Image plane |
| $\overline{\theta}_0$ | Initial parameter vector |
| $\beta_{ij}$ | Interaction terms |
| $\theta_{\text{int}}$ | Intrinsic parameters |
| $\mu_i$ | Mean of the $i^{th}$ factor level |
| $y_{ij}$ | Measured response at *ij* |
| $\beta_0$ | Overall average |
| $\beta_r$ | Pitch |
| $\hat{y}_{ij}$ | Predicted response at *ij* |
| $\beta_{jj}$ | Quadratic terms |
| $\beta_i$ | Regression coefficient |
| $\alpha_r$ | Roll |
| $\alpha_s$ | Significant level used in hypothesis testing |
| $\xi$ | Specified accuracy criterion |
| $\alpha$ | Star or axial points |
| $\mu_B(T)$ | The mean of the pixels in the group 'Background' at threshold 'T' |

| | |
|---|---|
| $\mu_F(T)$ | The mean of the pixels in the group 'Foreground' at threshold 'T' |
| $\sigma_B^2(T)$ | The variance of the pixels in the group 'Background' at threshold 'T' |
| $\sigma_F^2(T)$ | The variance of the pixels in the group 'Foreground' at threshold 'T' |
| $n_B(T)$ | Total number of pixels in the group 'Background' at threshold 'T' |
| $n_F(T)$ | Total number of pixels in the group 'Foreground' at threshold 'T' |
| $\tau_i$ | Treatment effect *i* |
| $\sigma^2$ | Variance |
| $\hat{\delta}'$ | Vector of the distortion parameters used in the backward camera model |
| $\overline{\theta}$ | Vector representing the unknown parameters |
| $\gamma_r$ | Yaw |
| $a_i'$ | Angle made by the axis of the bottom spheres |
| $a_i$ | Angle made by the axis of the top spheres |
| $\sigma^2_{Between}(T)$ | Between-group variances at threshold 'T' |
| $\alpha_x$ | The tilt angle of the image plane during X-axis displacements |
| $\alpha_y$ | The tilt angle of the image plane during Y-axis displacements |
| $\sigma^2_{Within}(T)$ | Within-group variances at threshold 'T' |
| $(D_u,D_v)$ | Conversion factors |
| $(k_1,k_2)$ | Radial distortion coefficients |
| $(p_1,p_2)$ | Tangential distortion coefficients |
| $(u,v)$ | Image coordinates |
| $(u_c,v_c)$ | Corrected image coordinates |
| $(u_d,v_d)$ | Distorted image coordinates |
| $(X_C,Y_C,Z_C)$ | Camera coordinates |
| $(CX_I,CY_I)$ | Image coordinates in the computer coordinate system |
| $(X_W,Y_W,Z_W)$ | World coordinates |
| $(\alpha_r, \beta_r, \gamma_r)$ | Euler angles |

| | |
|---|---|
| 0-D | Zero dimensional |
| 1-D | One dimensional |
| 2-D | Two dimensional |
| 2FI | Two-Factor Interaction |
| 3-D | Three dimensional |
| ANOVA | Analysis of Variance |
| *Asp* | Aspect ratio |
| **C** | Camera frame |
| C.V. | Coefficient of Variation |
| CCD | Central Composite Design |
| CCDs | Charge Coupled Devices |
| $C_e$ | Covariance matrix |
| ${}^{C}P_w$ | Point in the camera coordinate system |
| CV | Computer Vision |
| DLT | Direct Linear Transformation |
| DOE | Design of Experiments |
| $f$ | Effective focal length |
| $F$ | Function vector |
| F | Perspective transformation matrix |
| GCC | Geometric Camera Calibration |
| GUI | Graphical User Interface |
| $H$ | Center to center distance between two spheres |
| $h(x,y)$ | Gaussian kernel |
| IO | Interior Orientation |
| ${}^{I}P_d$ | Distorted image point with respect to the image coordinate system in the computer |
| $J$ | Jacobian matrix |
| $k$ | Number of factors |
| LED | Light Emitting Diode |

| | |
|---|---|
| $L_I$ | Distance from the projective center to the approximate image center |
| $L_W$ | Distance between the projective center and the center of the lined up top sphere |
| $M$ | Magnitude of the sobel gradient image |
| $m$ | Number of column pixels |
| MATLAB | Matrix Laboratory |
| $M_{ij}$ | $(i,j)^{th}$ moment |
| MV | Machine Vision |
| $n$ | Number of row pixels |
| $O_C$ $(U_0, V_0)$ | Image center or principal point |
| PC | Personal Computer |
| $P_d$ | Distorted image point |
| PIC | Peripheral Interface Controller |
| PRESS | Prediction Error Sum of Squares |
| PTM | Perspective Transformation Matrix |
| Pu | 2-D image point |
| $P_u$ | Undistorted image point |
| $P_W$ | 3-D world point |
| $P_x$ | Pixel size in X direction |
| $P_y$ | Pixel size in Y direction |
| $R$ | Radius of the image of the sphere |
| **R** | Rotation matrix |
| RGB | Red Green Blue |
| RSM | Response Surface Methodology |
| $s$ | Scale factor |
| $t_x$ | Translation in x direction |
| $t_y$ | Translation in y direction |
| $t_z$ | Translation in z direction |
| USB | Universal Serial Bus |

| | |
|---|---|
| $W$ | Displacement of the sphere |
| $\mathbf{W}$ | World frame |
| ${}^{w}P_{w}$ | Point in the world coordinate system |
| $x_i$ | Variable representing factor $i$ |
| $y$ | Response variable |

# Chapter 1

# Introduction

Machine Vision is one of the leading technologies of the present era. Machine Vision (MV) is the study of methods and techniques whereby artificial vision systems can be constructed and usefully employed in practical applications (Davies, 2004). As such, it embraces both the science and engineering of vision. It is an application of computer vision and image processing to the manufacturing industry for improving product quality and safety, as well as enhancing process efficiency and operational safety (Batchelor and Whelan, 1997). Whereas Computer Vision (CV) is a branch of Computer Science, MV is an area of specialization within Systems Engineering that includes not only the study of software but also the hardware environment and image acquisition techniques needed to apply it. Therefore, Machine Vision is a subfield of engineering that incorporates computer science, electronic engineering (hardware and software), engineering mathematics, physics (optics and lighting), mechanical engineering, and industrial

automation. A MV system often consists of components like digital cameras, suitable light sources, computer networks to control other manufacturing equipment, embedded processors or PCs, actuators, sensors etc. Machine Vision is a relatively new technology and its applications are diverse and wide ranging. Areas of endeavor include:

- Industrial applications for manufacturing, control and automation, quality control, inspection etc.,
- Spaceflight,
- Remote sensing,
- Medical image analysis,
- Robot vision,
- Industrial metrology,
- Reverse engineering,
- Image fusion,
- Tracking,
- Automatic assembly of electronics/mechanical components etc.,
- Monitoring of agricultural production, botany, agronomy, ecology and forestry.

Machine Vision was initially employed for two dimensional (2-D) measurements and focused primarily on image analysis. Later, complex three dimensional (3-D) measurement systems were developed leading to a convergence of automation technologies including vision, robotics and metrology. One of the most common industrial applications of machine vision systems is visual inspection; e.g. in the inspection of manufactured goods such as semiconductor chips, automobiles, food,

pharmaceuticals and other demanding applications that require high-speed, high precision, high-magnification, 24-hour operation, contactless detection, and/or repeatability of measurements. A diagrammatic representation of a typical 3-D Machine Vision system is shown in Figure 1.1:

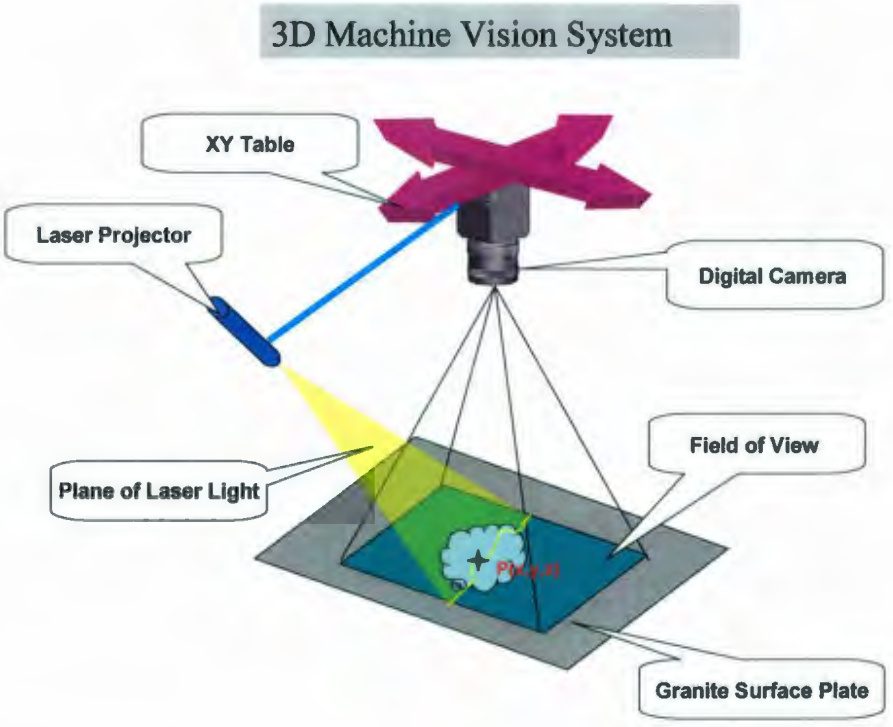



Figure 1.1: Diagrammatic representation of a 3-D machine vision system (Krouglicof, [Online])

Since the primary function of a Machine Vision system is to represent a real object by images, cameras are an indispensable component in a MV system. A camera must be calibrated before using it for high precision measurement applications. This important topic is covered in this research work.

Camera calibration has long been an important issue beginning with the photogrammetry community. With the increasing demand for high accuracy measurements, it has recently attracted researchers in the Computer Vision (CV) and robotic communities leading to a great demand for reliable and accurate calibration techniques for Machine Vision Systems. Compared with the high quality metric cameras used in photogrammetry, cameras commonly used in computer vision / machine vision have the following characteristics (Weng et al., 1990): (a) Image resolution is well-defined by spatial digitization and is relatively low (e.g. a typical Charge-Coupled-Device (CCD) sensing array has about $1024 \times 1024$ pixels); (b) Lenses used for video cameras are non-metric off-the-shelf lenses and exhibit a substantial amount of distortion; and (c) Camera assembly results in considerable internal misalignment. For example, the CCD sensing array may not be orthogonal to the optical axis of the lens and the center of the array may not coincide with the optical principal point; i.e. the intersection of the optical axis of the lens and the image plane. Due to these factors, camera calibration is an important consideration for high accuracy 3-D measurement systems. More details on camera calibration and the various techniques currently used for camera calibration are discussed in the following sections.

Camera calibration is an issue of fundamental importance in machine vision applications involving quantitative image analysis. In the field of machine vision, camera calibration refers to the experimental determination of a set of parameters which describe the image formation process for a given analytical model of the machine vision system. A complete set of calibration parameters includes both the intrinsic parameters that describe

the lens-camera-frame grabber combination as well as the extrinsic parameters that relate the position and orientation of the camera to a fixed reference frame. In other words, camera calibration is the process of determining the true parameters of the camera that produced a given video or still image. An accurate, reliable calibration procedure is essential for most industrial machine vision applications including mechanical metrology, robot assembly, reverse engineering, etc.

In 3-D Machine Vision, it is necessary to know the relationship between the 3-D object coordinates and the 2-D image coordinates (Heikkila, 2000). This transformation is determined in Geometric Camera Calibration (GCC) by experimentally determining the unknown parameters of the camera model. Typically camera calibration techniques are based on the observation of planar 2-D targets consisting of a large number of control points. In GCC, the assumption is made that the camera observes a set of features such as points or lines with known positions in some fixed world coordinate system (Forsyth and Ponce, 2003). In this context, camera calibration can be modeled as an optimization process where the discrepancy between the observed image features and their theoretical positions as predicted by the perspective transformation equations is minimized with respect to the camera's intrinsic and extrinsic parameters. Several linear and nonlinear approaches based on least square techniques can be used for solving this type of optimization problem. In general, once a camera has been calibrated it is possible to obtain accurate measurements from digitized images.

Camera calibration techniques were initially developed in the field of photogrammetry where reliable information about physical objects and the environment

are obtained through the process of recording, measuring, and interpreting photographic images and patterns of electromagnetic radiant energy and other phenomena for aerial imaging, remote sensing and surveying (Gruen and Huang, 2001). Later, new application areas including robot vision and industrial metrology emerged where camera calibration plays a vital role. The precision of the parameter estimation is an important factor for such applications. In general, camera calibration plays a crucial role in fields where high accuracy is expected. Such areas include industrial metrology, 3-D computer vision, robot vision, dimensional measurement of mechanical parts, automatic assembly of mechanical or electronics components, tracking, trajectory analysis, image fusion, reverse engineering, stereo vision and object localization in robotic tasks.

When a camera is used in a Machine Vision application, light from the environment is focused on an image plane and captured. This process reduces the dimensions of the data taken in by the camera from three to two i.e. light from a 3-D scene is stored on a 2-D image (Hartley and Zisserman, 2003). Each pixel on the image plane therefore corresponds to a ray of light from the original scene. Camera calibration (also called camera resectioning) determines which incoming light is associated with each pixel on the resulting image. In an ideal pinhole camera, a simple projection matrix is sufficient to do this. With more complex camera systems, errors resulting from lens misalignment and deformations in their structures can result in more complex distortions in the final image. The transformation from 3-D scene to 2-D image can be represented by a series of transformations; e.g. a lens distortion model, a projection matrix describing the lens/sensor assembly, a 3x3 rotation matrix, and a translation vector. More details on

the pinhole camera model are presented in Chapter 2. The camera projection matrix can be used to associate points in a camera's image space with locations in 3-D world space. Camera calibration is often used in stereo vision applications where the camera projection matrices of two cameras are used to calculate the 3-D world coordinates of a point viewed by both cameras.

The importance of camera calibration is described by Tsai (Tsai, 1987). Camera calibration in the context of 3-D machine vision is the process of determining the internal camera geometric and optical characteristics (intrinsic parameters) and/or the 3-D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters), for the following purposes:

1. *Inferring 3-D information from image coordinates*:

(a) The first is 3-D information concerning the location of the object, target, or feature. The applications include mechanical part dimensional measurement, automatic assembly of mechanical or electronics components, tracking, robot calibration and trajectory analysis. In the above applications, the camera calibration needs to be done only once.

(b) The second kind is 3-D information concerning the position and orientation of a moving camera (e.g. a camera held by a robot) relative to the target world coordinate system. The applications include robot calibration camera-on-robot configuration, and robot vehicle guidance.

2. *Inferring 2-D image coordinates from 3-D information*:

In model-driven inspection or assembly applications using machine vision, a hypothesis of the state of the world can be verified or confirmed by observing if the image

coordinates of the object conform to the hypothesis. In doing so, it is necessary to have both the intrinsic and extrinsic camera model parameters calibrated so that the 2-D image coordinate can be properly predicted given the hypothetical 3-D location of the object.

The above purposes can be best served if the following criteria for the camera calibration technique are met.

(a) *Autonomous*: The calibration procedure should not require operator intervention such as giving initial guesses for certain parameters, or choosing certain system parameters manually.

(b) *Accurate*: The camera calibration technique should have the potential of meeting the required accuracy. For example, applications such as mechanical part inspection and assembly require high accuracy, i.e. one part in a few thousandths of the working range.

(c) *Reasonably efficient*: The complete camera calibration procedure should avoid high dimensional nonlinear searches.

(d) *Versatile*: The calibration technique should operate uniformly and autonomously for a wide range of accuracy requirements, optical set-ups, and applications.

(e) *Need only common off-the-self camera and lens*: The advantages of using off-the-shelf camera and lens are versatility, high availability, user familiarity and low cost.

The goal of camera calibration is to correct the image displacements which occur due to elements of the camera's interior orientation (Computer Vision, [Online]). There are two general approaches used for camera calibration:

1) Model-based approaches: In a model based approach, a few predominant factors contributing to error are identified, and the errors are modeled using a mathematical

equation. Once models are generated for each contributing factor, the next step is to determine the values of the coefficients which best model the observed error. This can be accomplished in one of two ways:

a) Explicit Approach: In this approach, targets with known 3-D and 2-D coordinates or angular positions are used. Precision calibration frames, field approaches, etc. can be used to generate known positional targets. The pro and con of this approach are:

Pro: Precise

Con: Expensive and time consuming

b) Implicit Approach: In this approach, objects with known geometrical properties but with no known positional or angular orientation are used, e.g. Checkerboard approach, plumb-line approach. The pros and con are:

Pros:

- Relatively simple, fast and cheap
- Simple models can usually account for most of the error

Con: Not as precise as explicit approach

Therefore, the advantages and disadvantages of model-based approaches can be summarized as follows:

Pros:

- Relatively simple, fast and cheap
- Only a few factors, mostly optical configuration and lens distortion influence geometry

- Simple models account for most of the error

Cons:

- Can only remove errors represented by terms in the model
- Unknown causes of error are ignored

2) Mapping-based Approaches: With a mapping-based approach, no attempt is made to understand the individual contributing causes of error, rather the entire focus is on generating a comprehensive reality-to-image or image-to-reality mapping function. The advantages and disadvantages of mapping-based approaches can be summarized as follows:

Pros:

- Can handle all types of distortions
- Very precise, with subpixel accuracy
- Only needs to be done once
- Lends itself to automated approaches

Cons:

- Lot of explicit control is required
- Expensive and time consuming

The above paragraphs clearly answer questions like "What is the role of a camera in computer / machine vision applications?", "Why is camera calibration important in such vision based applications?" and "What are the different general approaches involved in camera calibration". In the following section a literature survey involving a more

detailed description and classification of the existing methods, models and techniques used for camera calibration in the photogrammetry and CV / MV communities will be presented.

## 1.1 Camera Calibration: Existing Methods, Models and Techniques

A camera calibration procedure is necessary in applications were metric information of the scene or environment is needed from 2-D images. Salvi et al. (2002) explain camera calibration by dividing the calibration procedure into two phases. Firstly, the camera is mathematically modeled by approximating the physical and optical behavior of the lens and image sensor. Secondly, the model parameters are estimated by using direct or iterative methods. Two kinds of parameters are considered in the camera model. *Intrinsic parameters* model the camera's internal geometry and optical characteristics. *Extrinsic parameters* relate the three dimensional position and orientation of the camera frame to a fixed world coordinate system which in turn provides metric information with respect to a user fixed coordinate system instead of the camera coordinate system. This section presents some of the current approaches adopted for camera calibration in close-range photogrammetry and computer vision. Various techniques and algorithms concerning camera calibration have been reported over the years in the photogrammetry and CV literature. The techniques proposed in the literature include those using 3-D calibration models (two or three orthogonal planes or a plane undergoing a pure translation, etc.), 2-

D objects (planar patterns undergoing unknown motions), 1-D objects (points aligned on a line) and 0-D features (self-calibration using unknown scene points) for calibration purposes (Zhang, 2002).

Much of the initial research work in camera calibration has been reported by the photogrammetry community (a few citations are: Brown, 1971; Faig, 1975; Fraser, 1997; Clarke and Fryer, 1998; Fraser, 2001; Remondino and Fraser, 2006 etc.) and by researchers working in the area of Computer Vision (a few citations are: Ganapathy, 1984; Faugeras and Toscani, 1987, 1986; Tsai, 1987; Weng et al., 1992; Wei and Ma, 1993, 1994; Heikkila, 2000; Zhang, 2008 etc.). The *Manual of Photogrammetry* (Slama, 1980) cites many research papers on camera calibration in the field of photogrammetry starting from the 1970's. These papers are categorized as close-range and terrestrial photogrammetric systems, architectural photogrammetry, industrial photogrammetry, underwater photogrammetry, etc. An overview of photogrammetric camera calibration techniques has also been reported in (Remondino and Fraser, 2006). Different camera models have been formulated and used in close-range photogrammetry, but generally sensor orientation and calibration is performed based on a perspective geometrical model by means of the bundle adjustment (Brown, 1971, 1974). A review of methods and models of the last 50 years is provided in (Clarke and Fryer, 1998). A mathematical model of the self-calibrating bundle adjustment with the extension of the collinearity model was formulated by (Fraser, 1997; Gruen and Beyer, 2001). The nonlinear collinearity equations include additional parameters such as principle distance and principle point offset as well as the radial and decentering distortion coefficients.

Various criteria for classifying calibration techniques are presented in the literature. Zhang classified the calibration techniques in computer vision and photogrammetry roughly into three categories; namely 3-D reference object based calibration, 2-D plane based calibration, and self calibration (Zhang, 2002). In 3-D reference object based calibration, camera calibration is performed by observing a calibration object whose geometry in 3-D space is known with very good precision. The object usually consists of two or three planes which are orthogonal to each other. Sometimes a planar pattern undergoing a precisely known translation is used to provide the equivalent of 3-D reference points (Tsai, 1987). It is stated that although these approaches require an expensive calibration apparatus and an elaborate set up, calibration can be done efficiently. Techniques involving 2-D plane based calibration require multiple observation of a single planar pattern presented at several different orientations (Zhang, 2008). In contrast to Tsai's approach, this technique does not require an expensive calibration apparatus and the calibration set-up is considered to be easier to implement since anyone can make such a 2-D calibration pattern (Zhang 2002). Techniques involving self-calibration in photogrammetry and CV do not make use of any calibration object and are considered to be 0-D approaches. Self-calibration is also known by the term auto-calibration where the metric properties of the camera and of the imaged scene are recovered from a set of uncalibrated images using constraints on the camera parameters or on the imaged scene (Remondino and Fraser, 2006). Generally three types of constraints are applied to perform self-calibration: scene constraints, camera motion constraints, or constraints on the camera intrinsic parameters. In general, a camera

undergoing a single displacement in a static (rigid) scene provides two constraints (Maybank and Faugeras, 1992; Zhang, 2008) on the camera's internal parameters using image information alone. Therefore, if images are taken by a single camera with fixed internal parameters, correspondences between three images are sufficient to recover both the internal and external parameters which allow the reconstruction of 3-D structure up to a similarity (Luong and Faugeras, 1997; Hartley, 1994). Auto-calibration matches corresponding features in multiple uncalibrated scene views (Faugeras et al., 1992; Hartley and Zisserman, 2000). Although these approaches are flexible, it is difficult to obtain reliable results because of the number of parameters that must be estimated (Zhang, 2008). Due to difficulty in initialization, auto-calibration results tend to be unstable (Sun and Cooperstock, 2005).

Various algorithms described in the photogrammetry and CV literature are generally based on perspective or projective camera models, with the most popular approach being the well-known self calibrating bundle adjustment, which was first introduced to close-range photogrammetry in the early 1970s (Remondino and Fraser, 2006). The literature mentions that depending on the nature of the application and the required accuracy, two basic functional models can be adopted for calibration purpose, namely a camera model based on perspective projection and a projective camera model supporting projective rather than Euclidean scene reconstruction.

Several different criteria can be applied to differentiate / classify camera calibration techniques (Salvi et al., 2002).

(1) Linear versus nonlinear camera calibration which is usually differentiated based on the modelling of lens distortion (Ito, 1991). The calibration results vary with the influence of lens distortion (Sun and Cooperstock, 2005). Tsai used a second order radial distortion model (Tsai, 1987) while Zhang employed both the second and fourth order terms (Zhang, 2000). Heikkila in his approach used a nonlinear fourth order radial distortion model and a second order tangential (decentering) distortion model (Heikkila, 2000). Lavest et al. (1998) used a sixth order radial distortion component, while Weng et al. (1992) introduced thin prism distortion which can be considered as a decentering distortion.

(2) Intrinsic versus extrinsic camera calibration: Intrinsic calibration deals with the estimation of the geometrical and optical parameters of the camera (Lenz and Tsai, 1988; Penna, 1991) whereas extrinsic calibration deals with the estimation of metric information concerned with the position and orientation of the camera in the scene (Liu et al., 1990; Wang, 1992).

(3) Implicit versus explicit models: Implicit calibration is the process of calibrating a camera without explicitly computing its physical parameters and is useful for both three-dimensional (3-D) measurement and generation of image coordinates (Wei and De Ma, 1994). Although implicit calibration techniques achieve high accuracy, they are computationally expensive and do not reveal the physical camera parameters.

(4) Methods using 3-D points (3-D world coordinates) as the calibration pattern (Heikkila, 2000; Tsai, 1987; Faugeras and Toscani, 1986; Hong and Yang, 1993) versus 2-D planar point arrays (Triggs, 1998; Tsai, 1987; Zhang, 2008).

(5) Point-based versus line-based methods (Caprile and Torre, 1990; Fryer and Brown., 1986): Point-based methods are more popular in photogrammetry with the only line-based approach of note being plumbline calibration which yields lens distortion parameters but not of interior orientation (Remondino and Fraser, 2006). Line-based methods are also used in the CV community (Salvi et al., 2002). Examples include calibration methods that use geometrical properties (also called geometric invariants) in the scene such as parallel lines (Echigo, 1989) or vanishing lines (Wang and Tsai, 1991) or other line features (Chen and Tsai, 1990).

A more specific classification of the different approaches can be made according to the calibration method used for parameter estimation of the camera model and the optimization technique employed:

(1) *Direct nonlinear optimization techniques*: In this category, equations are established that relate the parameters to be estimated with the 3-D coordinates of control points and their image plane projections (Weng et al., 1990). A calibrating technique is said to be nonlinear if the camera model includes any kind of lens imperfection. The objective of this type of optimization technique is to estimate the camera parameters iteratively by minimizing an objective function; i.e. to minimize residual errors of some equations. In photogrammetry the basis of the self calibrating bundle adjustment is an extended collinearity equation model which is a nonlinear technique (Remondino and Fraser, 2006). Most of the classical approaches in photogrammetry belong to this category (a few citations are: Brown, 1966; Abdel-Aziz and Karara, 1971; Wong, 1975; Faig, 1975; Slama, 1980). For instance, a rigorous and accurate modelling of the camera IO and lens

distortion parameters is provided through an iterative least-squares estimation process (Brown, 1971; Remondino and Fraser, 2006). The advantage of using this type of technique is that the camera model can be very general to cover many types of distortion (Weng et al., 1990) and is easily adapted to any arbitrarily accurate yet complex imaging model (Tsai, 1987). If the estimated model is good and correct convergence is reached, the algorithm can result in high accuracy (Weng et al., 1990). However, being an iterative algorithm there are certain drawbacks to this type of technique: (a) It requires a good initial guess to start the nonlinear search. If not, the procedure may end up with a bad or false solution. This violates the principle of automation. (b) It requires a computationally expensive, full-scale, nonlinear search (Tsai, 1987). (c) If distortion parameters are included in the parameter space of the camera model, the minimization may be unstable. The interaction between the distortion parameters and other camera parameters can lead to divergence or incorrect solutions (Weng et al., 1990).

(2) *Closed-form solution (Linear techniques)*: This type of scheme uses linear techniques to compute the transformation matrix relating 3-D points in the world coordinate frame to their 2-D projections in the image plane. The well known DLT (Direct Linear Transformation) method developed by Abdel-Aziz & Karara exemplifies such a technique. Linear techniques make use of the least-square method to obtain the transformation matrices. Parameters are computed directly through a non-iterative algorithm leading to a closed form solution (e.g. Abdel-Aziz and Karara, 1971; Ganapathy, 1984; Wong, 1975; Faugeras and Toscani, 1986; Weng et al., 1990). By first solving a set of linear equations, the algorithm computes a set of intermediate parameters

that can be expressed in terms of the desired camera parameters. The camera parameters are then determined from the intermediate parameters. Some of the advantages of this technique are: no nonlinear optimization is needed and hence the algorithm is computationally fast and simple providing a rapid calibration; and the camera model is simple without any nonlinear lens imperfections included. However, there are some drawbacks associated with this technique. First, lens distortion cannot be incorporated into linear techniques, resulting in low accuracy. Although the DLT method can be extended to incorporate distortion parameters (e.g. Shih et al., 1993), the corresponding formulation is not exact. The second disadvantage of linear methods is that in order to create a non-iterative algorithm, the actual constraints in the intermediate parameters are not considered. Consequently, in the presence of noise, the intermediate solution does not satisfy the constraints and the accuracy of the final solution is relatively poor (Weng et al., 1990). Moreover, it is sometimes difficult to extract some of the physical camera parameters from the transformation matrix due to the use of implicit camera calibration (Salvi et al., 2002). Some of the references cited for linear techniques can be found in (Faugeras and Toscani, 1986; Ganapathy, 1984; Wong, 1975; Hall et al., 1982; Ito,1991; Abdel-Aziz and Karara, 1971; Sutherland, 1974).

(3) *Two-step techniques*: Two-step techniques are a combination of linear and nonlinear techniques and make use of the advantages of the techniques previously described. As a first step, linear techniques are employed to compute or recover initial approximations for some of the parameters and as a second step the orientation and refined calibration parameters are computed iteratively. Some references related to two-step methods are

(Tsai, 1987; Lenz and Tsai, 1988; Wei and De Ma., 1994; Heikkila and Silven, 1997). The advantages of their methods are: (a) They permit a rapid calibration by considerably reducing the number of iterations required and the number of parameters to be estimated iteratively; (b) A closed-form solution can be derived for a major set of parameters initially (Weng et al., 1990); and (c) The convergence is nearly guaranteed because the linear approximation of the parameters obtained in the first step is in the vicinity of the final solution (Salvi et al., 2002).

Geometric Camera Calibration in CV and MV make use of 3-D or planar 2-D calibration models. These models typically consist of multiple circular control points or checkerboard patterns where the vertices of the squares serve as control points. Commonly adopted methods in CV/MV are those of Tsai (1987), Zhang (2008) and Heikkila (2000). These are the most widely used plane-based calibration algorithms. All are based on the pinhole camera model including terms for modelling lens distortions. The various methods differ in their implementation, formulation and mathematical modeling.

Tsai's method (Tsai, 1987) is an example of a two-step technique that consists of a combination of linear and nonlinear methods. A technique for three-dimensional camera calibration for machine vision based metrology using off-the-shelf TV cameras and lenses is described in Tsai (1987). The proposed two-stage technique is aimed at efficiently computing the camera external position and orientation relative to the object reference coordinate system as well as the effective focal length, radial lens distortion, and image scanning parameters. The calibration model assumes that some camera

parameters are provided by the manufacturer hence no initial approximation of these parameters is required, which makes the technique fast. The approach is based on the radial alignment constraint which is a function of five of the six extrinsic camera parameters; namely, the relative rotation and translation between the camera and the calibration points with the exception of the z translational vector. Hence, by using this radial alignment constraint (parallelism constraint) all of the extrinsic parameters except for $t_z$ are computed in the first stage. This constraint considerably reduces the dimensionality of the unknown parameter space while solving a subset of calibration parameters (coefficients of the homogeneous transformation matrix) with linear equations. In the second stage, all other parameters are estimated by nonlinear optimization. The intrinsic parameters are computed using normal projective equations. A second order radial distortion model is used while no decentering or tangential distortion effects are considered. The optimization technique does not use the full camera model in order to speed up the performance (Zollner and Sablatnig, [Online]). Moreover, the number of iterations is considerably reduced by the proposed algorithm although all the parameters are iteratively optimized. This method works with a single plane or multiple plane images and with 2-D or 3-D calibration grids as targets. In the case of multiple calibration images, a planar pattern is moved to different levels by a $z$ stage.

Zhang's method (Zhang, 2008) uses 2-D planar targets (e.g. a checkerboard pattern) as a calibration target and uses the vertices of the squares as control points to compute a projective transformation. The technique requires the camera to observe the target presented at several different orientations producing $n$ different images (with $n \geq 2$)

either by moving the camera or the planar pattern. The proposed procedure consists of a closed-form solution for estimating the camera's intrinsic (interior) and extrinsic (exterior) parameters, and a linear least-squares solution for recovering the coefficients of radial distortion. This is followed by a nonlinear optimization technique (Levenberg-Marquardt Algorithm) based on the maximum likelihood criterion for refining all the recovered parameters by minimizing the reprojection error. Zhang's approach is similar to an earlier technique proposed by Triggs (1998) that requires a minimum of five views of a planar target.

The technique developed by Heikkila (Heikkila, 2000) consists of a forward camera model which converts the 3-D world coordinates to distorted image coordinates and a backward camera model which transforms the distorted camera coordinates to lines-of-sight in the 3-D world coordinate system, or to the intersection of these lines with a known 2-D plane. More on this camera model and technique is described in Chapters 2 and 3. The calibration procedure is mainly intended to be used with circular control points. The camera model includes eight intrinsic and six extrinsic parameters. The calibration model first extracts the initial estimates of the camera parameters using a closed-form solution based on a Direct Linear Transformation (DLT) solution and then a nonlinear least-squares estimation employing numerical techniques like the Levenberg-Marquardt algorithm to refine the parameters of the forward camera model. It makes use of distortion models and inverse distortion models to distort and correct the image coordinates respectively. The model considers two coefficients of radial and two

coefficients of tangential (decentering) distortion. The proposed method works with single or multiple images and with 2-D or 3-D calibration grids.

To summarize, chronologically Tsai's algorithm is the oldest technique that is still widely used in CV / MV, and there are numerous implementations in C / C++ or other high-level computer languages (Zollner and Sablatnig, [Online]). Zhang's method, which makes use of advanced concepts in projective geometry, offers significant advantages in terms of flexibility in comparison to other classical techniques which use expensive equipment such as two or three orthogonal planes. Heikkila's method uses an array of circular control points as a target and hence the method requires additional subpixel detection techniques and ellipse fitting to extract the circular control points from the synthetic images making the whole procedure more complicated. Still this method gives quite accurate results in estimating the camera parameters including lens distortion coefficients. A camera calibration toolbox for MATLAB has also been proposed by Heikkila (Heikkila, 2000).

## 1.2 Proposed Project

The quest for the best solution to the Geometric Camera Calibration has been a research problem for many years in the Computer / Machine Vision communities. One of the most systematic calibration procedures for 3-D machine vision applications was proposed by Heikkila (2000) in which a comprehensive set of camera parameters is automatically evaluated by observing a calibration target consisting of two perpendicular planes, each

with 256 circular control points (Figure 1.2). Other similar techniques employ a checkerboard pattern as a target and use the vertices of the squares as control points (Figure 1.3). While these techniques are sound from a theoretical point of view, they do not adequately speak to the question of measurement accuracy. The following figures represent two types of calibration targets used mainly for machine vision applications as explained in the previous sections.

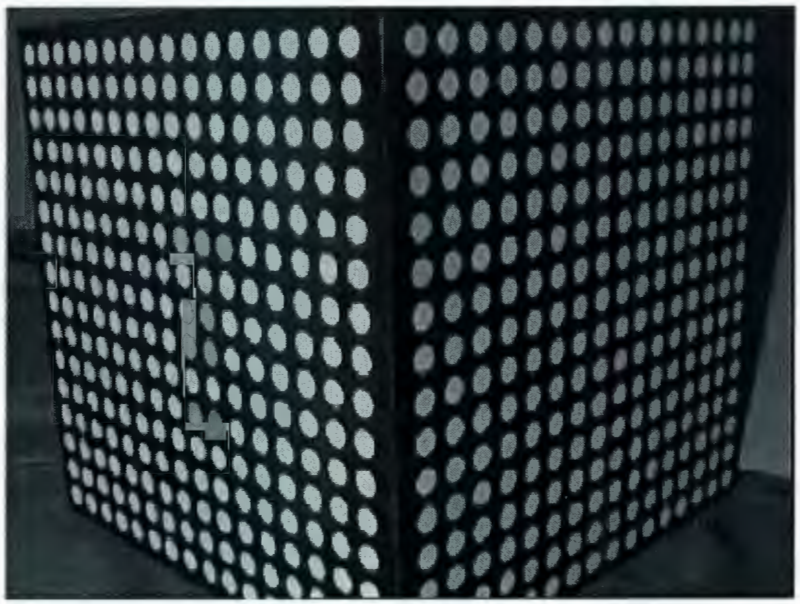


Figure 1.2: 3-D Calibration target (Heikkila, 2000)

Figure 1.3: Checkerboard pattern as target
(http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/calib_example/index.html)

This study addresses the problems associated with geometric camera calibration as well as the question of accuracy. The broad use of camera calibration in the field of machine vision and computer vision is the primary motivation for understanding the major factors that influence the accuracy of camera calibration as well as the accuracy that can be achieved within the practical limits of most research environments. Heikkila's camera calibration technique was selected for the detailed investigation in this study.

### 1.2.1 Methodology Overview

The study starts with the implementation of the technique and algorithms proposed by Heikkila for geometric camera calibration. The performance of the technique is then evaluated by analyzing a virtual camera image of a hypothetical calibration target that is created by applying a set of known camera parameters. The entire performance evaluation process can be briefly described as follows:

Step 1: Define a 3-D target which consists of an array or arrays of control points whose object coordinates in the world coordinate system are known.

Step 2: By applying a set of known camera parameters, calculate the image coordinates of these control points in the image coordinate frame with the Euclidian and projective transformations as well as with lens distortion.

Step 3: Perform the calibration by employing the proposed calibration method in order to extract the camera's intrinsic and extrinsic parameters and compare with the known values used in Step 2.

As an initial start to the proposed research, Heikkila's calibration algorithm was verified with a set of synthetic images of a simulated 3-D target. A set of intrinsic and extrinsic parameters were computed from the synthetic images by applying Heikkila's calibration algorithm. These parameters were found to be exactly equal to the values used to generate the synthetic image which implies that the algorithm works properly in estimating the parameters. The camera modelling in GCC is broken down into 4 steps as described in Figure 1.4:

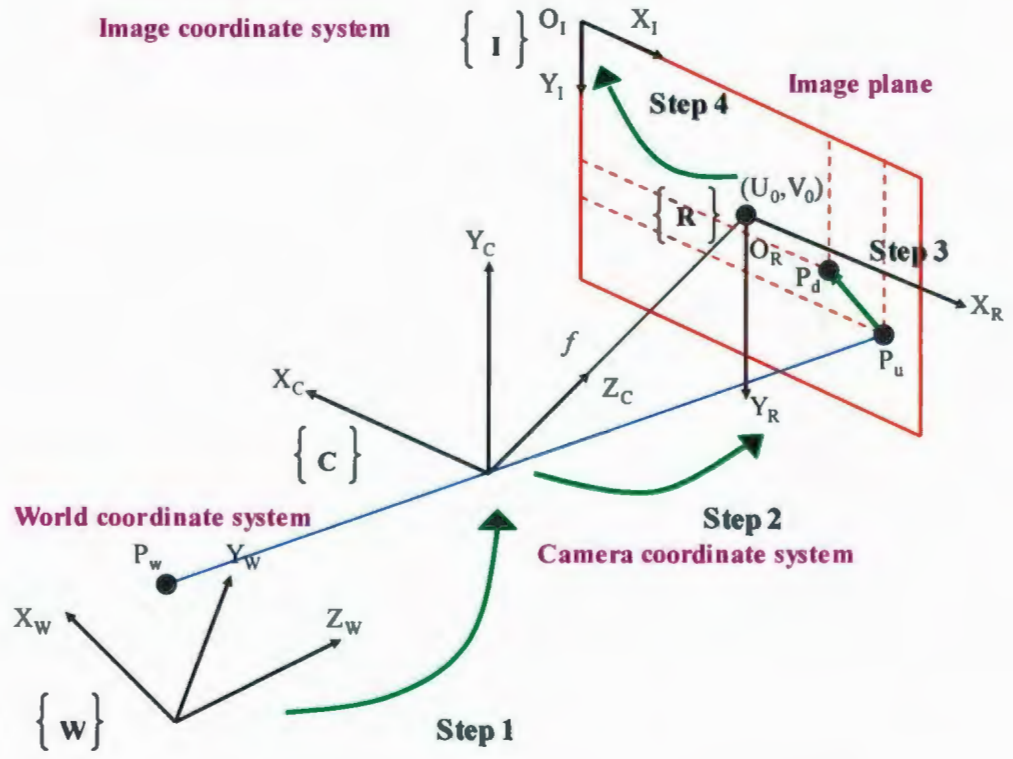



Figure 1.4: The geometric relation between a 3-D object point and its 2-D image projection

Step 1: Consists of the transformation of a point ${}^{W}P_w$ in the world coordinate system to the camera coordinate system obtaining ${}^{C}P_w$. This transformation is done by using rotation matrix and a translational vector.

Step 2: Consists of a projective transformation of the point ${}^{C}P_w$ to the image plane obtaining the point $P_u$.

Step 3: Models the lens distortion which transforms the point $P_u$ to $P_d$.

Step 4: Consists of the coordinate system transformation of the point $P_d$ to the image coordinate system of the computer producing ${}^{I}P_d$ which is measured in pixels.

More on camera modeling and the mathematical relationship between the transformations involved in this study are discussed in Chapter 2.

Camera calibration has been studied extensively in Computer / Machine vision, and various techniques have been employed; however, it can be said that while these techniques are sound from a theoretical point of view, they do not adequately speak to the question of measurement accuracy. Hence the study was extended to gain an understanding of how much accuracy is needed on the target points for an optimal calibration. This study addresses the following fundamental question: Assuming a certain tolerance or uncertainty in the calibration target, what is the expected error with respect to the measured camera parameters and what is the impact on the final 3-D machine vision application? Later, the study on the accuracy analysis was extended to understand the problems associated with geometric camera calibration through the application of Design of Experiments (DOE). A response surface methodology (RSM), namely a Central Composite Design (CCD), is carried out for this purpose. The objective of this study is to find out the significant factors leading to inaccurate results during the process of camera calibration. Based on the results of this study, the final focus was to propose a new approach towards camera calibration in which image center and focal length are calculated independently of the lens distortion. The experimental setup includes an X-Y table with camera, LED lighting, stepper motor driven ball screws for controlling the table motion, a microcontroller interface circuit and a computer. The software, which is developed in MATLAB, consists of a series of image processing algorithms to extract the image points with subpixel accuracy.

### 1.2.2 Significance of the study

Geometric camera calibration introduces the analytical tools necessary to establish the quantitative constraints between the image measurements and the position and orientation of geometric figures measured in some arbitrary coordinate system. Once a camera has been calibrated, it is possible to associate with an image point a well defined ray which passes through this point and the camera's optical center as well as to perform three dimensional measurements from a digitized picture (Forsyth and Ponce, 2003). An accurate, reliable calibration procedure is essential for most industrial machine vision applications including areas like mechanical metrology, robot assembly, reverse engineering, etc. An optimal calibration technique should produce unbiased and minimum variance estimates of the camera parameters. In practice, this is quite difficult to achieve due to various error sources affecting the imaging process. The precision of the parameter estimates is an important aspect in all camera calibration techniques. This study focuses mainly on the role of "accuracy" in geometric camera calibration by identifying a few predominant factors contributing to error.

## 1.3 Document Organization

This document is organized as follows: Chapter 2 describes the camera model in general with a detailed explanation of the pinhole camera model, perspective geometry and other

mathematical models of the image formation process. Chapter 3 describes the relevant background and motivation which includes an overview of Heikkila's calibration techniques and algorithms. Chapter 4 describes the work done for the implementation and validation of Heikkila's method with a discussion on the accuracy analysis of the calibration model, including methodology and results. Chapter 5 describes the application of Design of Experiments (DOE) to analyze the problems associated with geometric camera calibration. Included is a description of DOE, experimental design considered, a description of the methodology and a discussion of the results obtained. Chapter 6 describes the new approach used for camera calibration with a description of the technique as well as experimental results acquired with the proposed technique. Finally, in Chapter 7, results are summarized, conclusions are drawn, contributions are highlighted and recommendations on future work are presented.

# Chapter 2

# Camera Model

In machine vision we need to know aspects of a camera's image-formation process that range from simple properties, such as magnification and focussed distance, to more complex image properties such as perspective projection and image defocus (Gruen and Huang, 2001). In order to have computationally efficient, closed-form equations for the more complex properties, we use models that are based on simplifications or abstractions of the true image-formation process of the lens. The two most commonly used abstract models are the pinhole camera model and the thin-lens model. In this research work of camera calibration, a pin-hole camera model, which is the simplest kind of camera model, is demonstrated for the mathematical modeling of image formation process. Models that account for lens distortion first try to move an observed image point to another point corresponding to the pinhole camera model (Tsai, 1987; Heikkila, 2000). A camera model is an algorithm for calculating a line where an observed point should lie, given its

coordinates in the image (Ryberg et al., 2006). Or conversely, knowing the relative position and orientation of the camera and an object point, calculate where it would end up in the image.

Depending on the accuracy requirements, the camera model is based on either orthographic or perspective projection (Heikkila, 2000). Orthographic transformation, which is the roughest approximation, assumes that the objects in 3-D space are orthogonally projected on the image plane. It is more suitable for vision applications where the requirements of the geometric accuracy are somewhat low. Due to its linear nature, it provides a simpler and computationally less expensive solution than perspective projection which is a nonlinear form of mapping. For 3-D motion estimation and reconstruction problems, perspective projection gives an idealized mathematical framework which is actually quite accurate for high quality camera systems. For off-the-self systems, the perspective projection model is often augmented with a lens distortion model. Analytical approaches to calibration have a common starting point, that is, the introduction of specific analytical models for radial and decentering (tangential) lens distortion directly into the projective equations (Slama, 1980).

## 2.1 Pinhole Camera Model: A Brief Introduction

The pinhole camera model describes the mathematical relationship between the 3-D (object) coordinates of a point in the world frame and its 2-D image coordinates in the image frame of an ideal pinhole camera; i.e. where the camera aperture is described as a

point (pinhole) and is used instead of lenses to focus light. A pinhole camera, also known as camera obscura (literally, dark chamber) can be imagined as a light-proof box with a small hole in one of its sides and a translucent plate on the opposite side. Light from a scene passes through this small hole and projects an inverted image on the translucent plate of the box. The inverted image is created by the perpective projection. Hence, it is convenient to consider a virtual image associated with a plane lying in front of the pinhole, at the same distance from it as the actual image plane, which will produce an image equivalent to the actual one (Forth and Ponce, 2003). Imaging devices such as cameras with small apertures, and even a human eye in bright light act like a pinhole camera. Figure 2.1 represents the principle of pinhole camera, where light rays from an object pass through a small hole to form an inverted image.

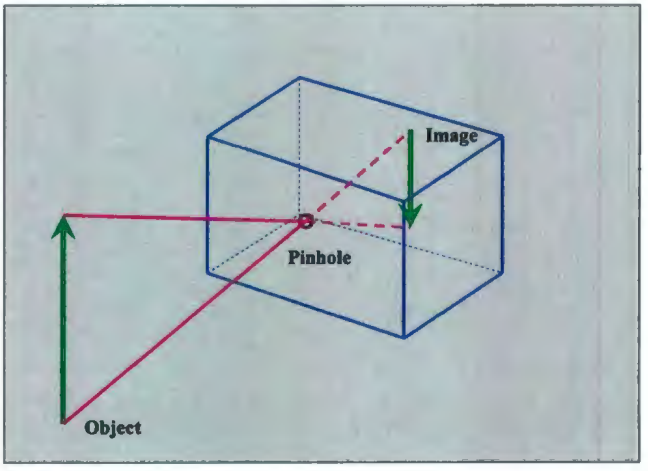



Figure 2.1: Principle of pinhole camera

The smaller the hole, the sharper the image is. But it reduces the amount of light reaching the image plane making the projected image appear to be dimmer. On the other hand, if the hole is large in size, the image looks brighter but gives blurry pictures. Optimally, the size of the aperture should be 1/100 or less of the distance between it and the screen. Pinhole cameras in combination with Charge Coupled Devices (CCDs) are often used for surveillance because they are difficult to detect.

An ideal pinhole camera model does not include secondary effects such as geometric / optical aberrations or blurring of unfocused objects caused by the lenses and finite or larger sized apertures. It also does not take into account the fact that most modern practical cameras have only discrete image coordinates (e.g. the commercially available off-the shelf cameras). Hence an ideal pinhole camera model can only be considered to be a linear model for mapping a 3-D scene and a 2-D image. If the goal is to get highly accurate 3-D geometric measurements, especially in machine vision applications, some of the lens distortion effects are compensated for by applying suitable coordinate transformations on the image coordinates and by adding a nonlinear lens distortion model to the existing pinhole camera model (more detailed descriptions of these mathematical formulations are explained in the following sections). Normally, other effects are negligible if a high quality camera is used. Such a pinhole camera model can generally describe how a camera depicts a 3-D scene. A pinhole camera model is widely used in various vision based applications involving areas like computer / machine vison, computer graphics, image processing, and pattern recognition.

## 2.2 The Geometry of Pinhole Camera Model

The pinhole camera model is based on the principle of collinearity, where each point in the world / object space is projected by a straight line through the projection center (the origin of the camera coordinate system) onto the image plane (Heikkila and Silven, 1997). Let us consider a pure perspective projection (i.e. pinhole) camera model illustrated in Figure 2.2 which often provides an acceptable approximation of the imaging process. The camera model assumes the laws of Gaussian optics with an extension for the aberration of geometric lens distortions (i.e. radial and tangential distortions). Hence, the camera model consists of a linear part corresponding to the ideal pinhole camera model (i.e. a first order approximation of the real camera projection) and a nonlinear part corresponding to the lens distortion model. The incorporation of nonlinear models is necessary to accurately model the lenses which are useful for applications requiring greater precision. The mathematics related to the geometry of a pinhole camera model is explained in the following sections:

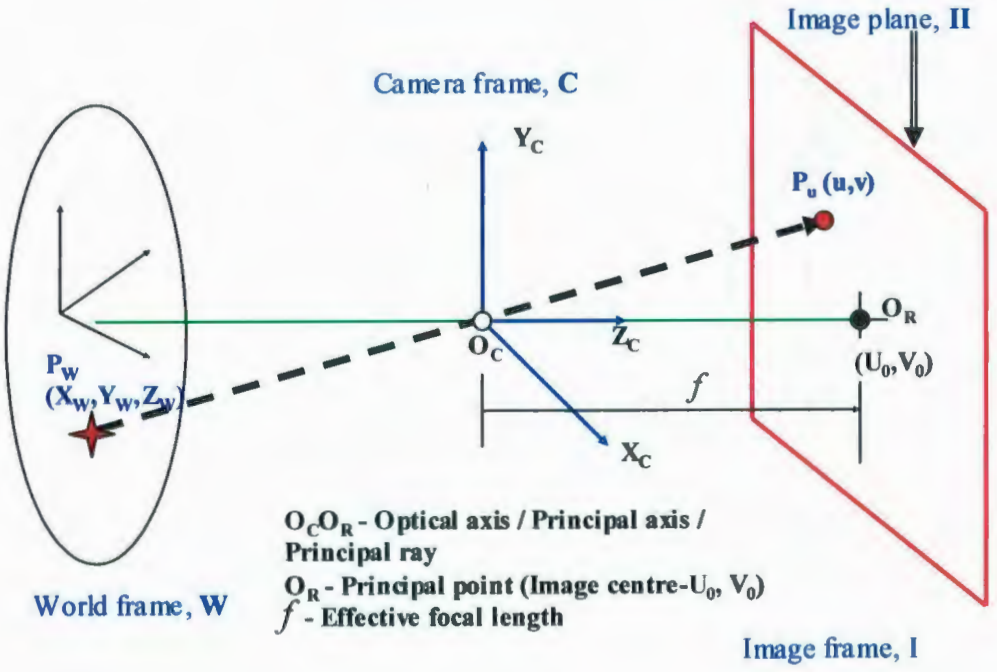



Figure 2.2: Ideal pinhole camera model

## 2.3 Definition of Various Coordinate Frames

1. *World Coordinates*

The coordinates of a point with respect to the object / world coordinate frame are denoted by ($X_W$, $Y_W$, $Z_W$). The geometry of the calibration target with respect to this coordinate frame is known a priori.

2. *Camera Coordinates*

The coordinates of a point with respect to the camera coordinate frame are denoted by ($X_C$, $Y_C$, $Z_C$). The origin of this coordinate frame $O_C$, coincides with the projection center

of the lens. The image plane, $\Pi$, is parallel to the $X_C Y_C$ plane and it is displaced a distance $f$ (i.e. the effective focal length) from $O_C$ along the $Z_C$ axis.

3. *Image Coordinates*

The two dimensional coordinates of a point on the image plane are denoted by (u,v). The origin of the coordinate frame $O_R$ $[(U_0, V_0)]$ lies at the intersection of the optical axis of the lens and the image plane. This point is known as the principal point or the image centre. Due to the lens distortions, the image points (u,v) are displaced. The distorted image points are denoted by $(u_d, v_d)$ and once the correction has been applied, these points are denoted by $(u_c, v_c)$.

Let '$P_W$' be an arbitrary 3-D point located in the world frame as can be seen in Figure 2.2. The projection of '$P_W$' corresponds to the point '$P_u$' in the image frame which is defined by the coordinates $[u,v]^T$. The coordinates of the point '$P_u$' can be determined from the homogeneous coordinates given by the following transformation:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \propto \begin{bmatrix} su \\ sv \\ s \end{bmatrix} = F \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = PM \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{2.1}$$

where F is the perspective transformation matrix (PTM) that can be defined as the product of two matrices P and M. The matrix P contains the intrinsic camera parameters:

$$P = \begin{bmatrix} Asp.f & 0 & U_O & 0 \\ 0 & f & V_O & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.2}$$

The parameter $s$ is the scale factor and *Asp* is the aspect ratio. The parameter $f$ and $U_0, V_0$ were previously defined as the effective focal length and image centre respectively. The matrix M is a 4 by 4 matrix describing the mapping from the world frame (**W**) to the camera frame (**C**), which contains the extrinsic camera parameters. It is composed of two parts as follows:

$$M = \begin{bmatrix} R^{3\times3} & t^{3\times1} \\ 0 & 1 \end{bmatrix} \tag{2.3}$$

where $t = [t_x \quad t_y \quad t_z]^T$ describes the translation between two frames (i.e. between **W** and **C**), and **R** is a 3 by 3 orthonormal rotation matrix that can be expressed in terms of the three Euler angles: $\alpha_r(roll), \beta_r(pitch), \gamma_r(yaw)$, that define a sequence of three elementary rotations around z, y, and x-axes, respectively. The rotations are performed first about z-axis by an angle $\alpha_r$, then about y-axis by an angle $\beta_r$, and finally about x-axis by an angle $\gamma_r$.

## 2.4 Camera Parameters

In practice, the world and camera coordinate system are related by a set of physical parameters such as the effective focal length of the lens, the position of the image center, the size of the pixels and the position and orientation of the camera. Generally, the

objective of the explicit camera calibration procedure is to determine the optimal values for these parameters based on image observations of a known 3-D target. In the case of self-calibration the 3-D coordinates of the target points are also included in the set of unknown parameters (Heikkila and Silven, 1997). These physical camera parameters are commonly divided into two, i.e. extrinsic and intrinsic parameters. These parameters are explained in the following paragraphs:

1. *Extrinsic Parameters (Exterior Projective Parameters)*

These parameters relate the camera's coordinate system to a fixed world coordinate system and define its position and orientation in space. In other words these parameters are required to define the transformation from 3-D object / world coordinate system to the 3-D camera coordinate system centered at the optical center. The extrinsic parameters also describe the relationship between cameras in multi-camera systems. The extrinsic parameters are the three Euler angles ($\alpha_r$, $\beta_r$, $\gamma_r$) and the three components of translation ($t_x$, $t_y$, $t_z$).The Euler angles can be computed by decomposing the rotation matrix **R** given as,

$$\textit{Rotation matrix, } \mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{2.4}$$

2. *Intrinsic Parameters (Interior Projective Parameters)*

The intrinsic parameters relate the camera coordinate system to the idealized image coordinate system, in which the image coordinates have their origin at the principal point.

In other words these parameters are required for the transformation from 3-D object / world coordinates measured with respect to the camera coordinate frame to 2-D image coordinates. The intrinsic parameters basically determine how light is projected through the lens onto the image plane of the sensor and models the internal geometry and optical characteristics of the camera. The intrinsic parameters are:

a. *Effective focal length ( $f$ )* – image plane to projective center distance.

b. *Image center* ($U_0, V_0$) – the point where the optical axis pierces the image plane, also called the principal point.

c. *Scale factor ( $s$ )* – results from the difference in frequency between the pixel clock of the camera and the acquisition rate of the frame grabber. This is only an issue in older machine vision systems that employ camera with an analog output signal. The scale factor is equal to 1 in modern digital cameras where the pixels are usually square.

d. *Lens distortion coefficients*: Actual cameras are not perfect and sustain a variety of optical aberrations (Weng et al., 1992). Off-the self consumer grade camera lenses exhibit inherent nonlinear distortion properties that must be accounted for in high accuracy machine vision applications. In this research two types of lens distortion are considered, radial and tangential / decentering distortions. Figure 2.3 illustrates radial and tangential distortion.

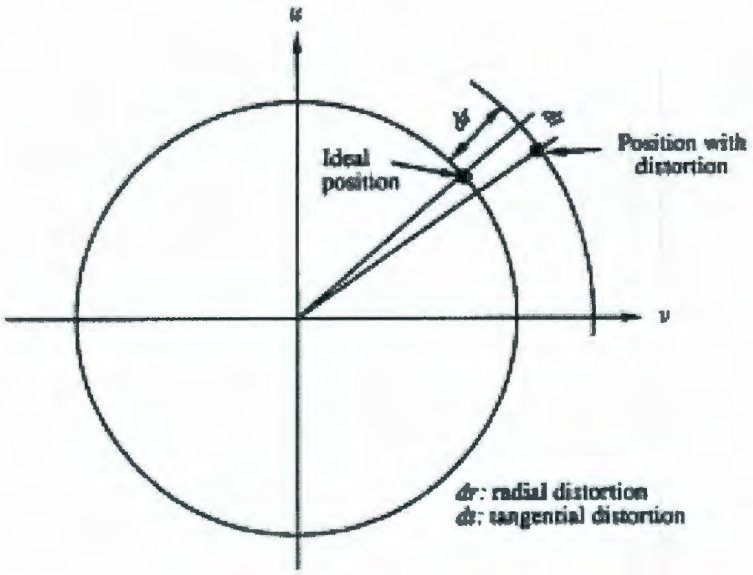



Figure 2.3: Radial and tangential distortions (Weng et al., 1992)

1. *Radial distortion coefficients ($k_1$, $k_2$)* – Radial distortion causes the actual image point to be displaced radially in the image plane; i.e. an inward or outward displacement of a given image point from its ideal location. Figure 2.4 illustrates the effects of radial distortion with respect to no distortion.

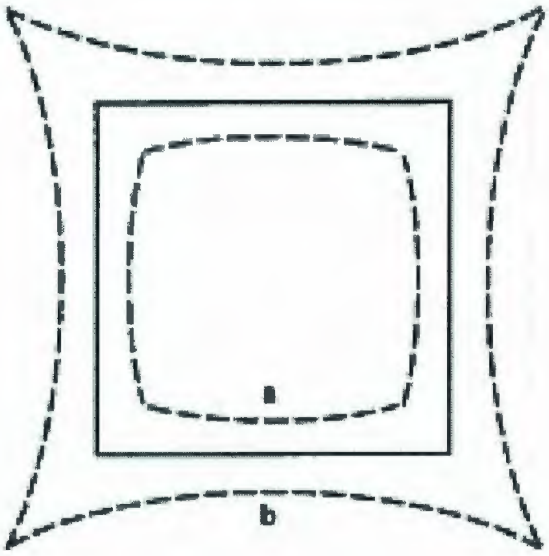



Figure 2.4: Effect of radial distortion. Solid lines: no distortion; dashed lines: with radial distortion (a: negative, b: positive) (Weng et al., 1992)

Radial distortion is a function of the distance separating the optical axis from the point of interest. This type of distortion is mainly caused by an incorrect radial curvature of the lens elements. A negative radial displacement of the image points is referred to as barrel distortion (Figure 2.5). It causes outer points to crowd together and scale to decrease (Weng et al., 1992). This leads to a decrease in the image magnification with the distance from the optical axis. Apparently, the effect is that of an image which has been mapped around a sphere.

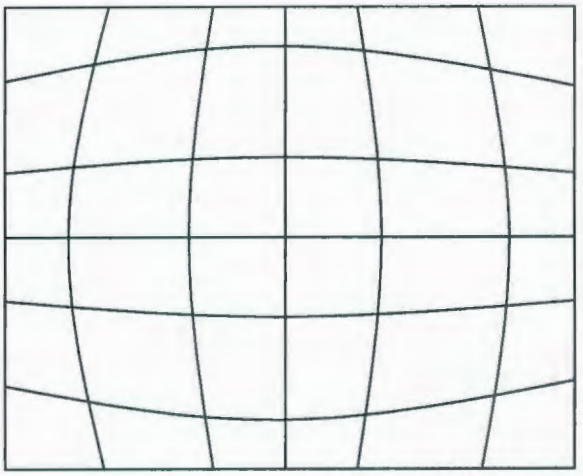


Figure 2.5: Barrel distortion

A positive radial displacement is referred to as pincushion distortion (Figure 2.6). It causes outer points to spread and the scale to increase. This leads to an increase in the image maginification with the distance from the optical axis. The apparent effect is that, lines that do not go through the centre of the image are bowed inwards towards the centre of the image. This type of distortion is strictly symmetric about the optical axis.

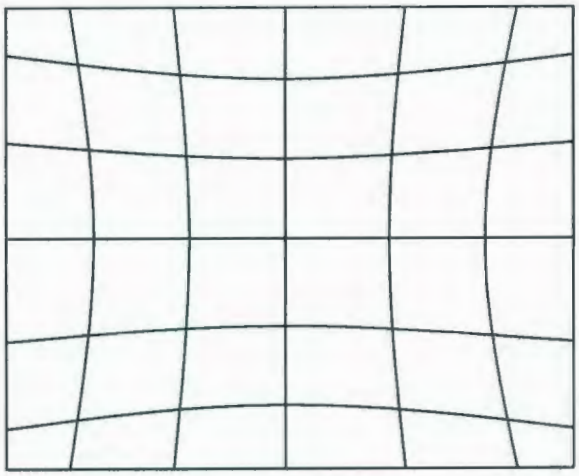


Figure 2.6: Pincushion distortion

2. *Tangential distortion coefficients ($p_1$, $p_2$)* - Tangential distortion is an image defect that occurs when the individual lens elements in a compound lens assembly are not strictly collinear. This misalignment of the optical centers causes the displacements of the image points to occur at right angles to a radius from the center of the field. In Figure 2.3, the displacement of the image point is perpendicular to the radial distortion indicated by *dt*. Figure 2.7 illustrates the effects of tangential distortion, showing the maximum and minimum axis of tangential distortion.

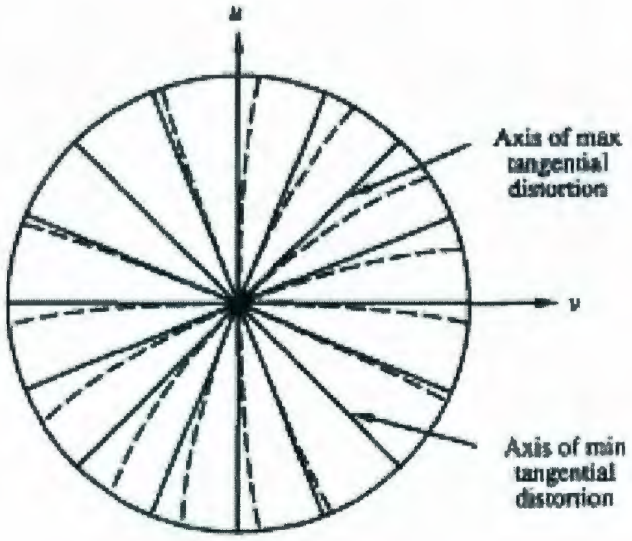



Figure 2.7: Effect of tangential distortion. Solid lines: no distortion; dashed lines: with tangential distortion (Weng et al., 1992)

Other distortion types have also been proposed in the literature (Heikkila and Silven, 1997). For example, Melen (1994) proposes a correction term for the linear distortion that occurs when the image axes are not orthogonal. In most cases this error is small and the distortion component is insignificant. Another error component is thin prism distortion. It arises from imperfect lens design and manufacturing, as well as camera assembly. This type of distortion can be adequately modeled by the adjunction of a thin prism to the optical system, causing additional amounts of radial and tangential distortions (Faig, 1975; Weng et al., 1992).

## 2.5 Mathematical Model of Image Formation Process (Camera Model)

### 2.5.1 Coordinate Transformations

The transformation describing the image formation process can be summarized as follows:

1. *World (i.e. Object) Frame to Camera Frame*

This transformation consists of a rotation followed by translation.

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix}^T = \begin{bmatrix} X_W \\ Y_W \\ Z_W \end{bmatrix}^T \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}^T \tag{2.5}$$

The parameters which must be identified are the nine rotation parameters and the three components of translation namely, $r_{11}, r_{12}, r_{13}, r_{21}, r_{22}, r_{23}, r_{31}, r_{32}, r_{33}, t_x, t_y, t_z$.

The nine rotation parameters can be expressed in terms of the three Euler angles, $\alpha_r, \beta_r, \gamma_r$ as follows (Craig, 1986):

$$
\begin{aligned}
r_{11} &= \cos\alpha_r \cos\beta_r \\
r_{12} &= -\sin\alpha_r \cos\gamma_r + \cos\alpha_r \sin\beta_r \sin\gamma_r \\
r_{13} &= \sin\alpha_r \sin\gamma_r + \cos\alpha_r \sin\beta_r \cos\gamma_r \\
r_{21} &= \sin\alpha_r \cos\beta_r \\
r_{22} &= \sin\alpha_r \sin\beta_r \sin\gamma_r + \cos\gamma_r \cos\alpha_r \\
r_{23} &= -\cos\alpha_r \sin\gamma_r + \sin\alpha_r \sin\beta_r \cos\gamma_r \\
r_{31} &= -\sin\beta_r \\
r_{32} &= \cos\beta_r \sin\gamma_r \\
r_{33} &= \cos\beta_r \cos\gamma_r
\end{aligned}
\tag{2.6}
$$

Conversely, the Euler angles can be computed from the elements of **R** by applying the following decomposition:

$$
\begin{aligned}
\alpha_r &= a\tan 2(r_{21}/\cos\beta_r, r_{11}/\cos\beta_r) \\
\beta_r &= a\tan 2(-r_{31}, \sqrt{r_{11}^{\,2} + r_{21}^{\,2}}) \\
\gamma_r &= a\tan 2(r_{32}/\cos\beta_r, r_{33}/\cos\beta_r)
\end{aligned}
\tag{2.7}
$$

The function $a\tan 2(x, y)$ is a two-argument inverse tangent function in MATLAB giving the angle in the range $[-\pi, \pi]$.

2. *Camera Frame to Sensor Frame (i.e. undistorted image coordinates measured with respect to the centre of the sensor)*

This transformation describes the single point perspective projection for an ideal lens-sensor assembly. The perspective projection equations are derived from the collinearity of the 3-D point '$P_W$', its image point '$P_u$', and the center of the pinhole '$O_C$' (Forsyth and Ponce, 2003). The coordinates of '$P_W$' in the world coordinate system are represented by $(X_W, Y_W, Z_W)$ and its projection in the image plane is represented by the point '$P_u$' with

coordinates $(u,v)$ (See Figure 2.2). Since the image plane is two dimensional and is parallel to the camera frame, the third coordinate of the image point located in the image plane can be considered to be the effective focal length '$f$'. Since the points '$P_W$', '$O_C$', and '$P_u$' are collinear, it can be written as $\overrightarrow{O_C P_W} = \lambda \overrightarrow{O_C P_u}$ for some number $\lambda$, so

$$\left.\begin{aligned} u &= \lambda\ X_W \\ v &= \lambda\ Y_W \\ f &= \lambda\ Z_W \end{aligned}\right\} \quad \lambda = \frac{u}{X_W} = \frac{v}{Y_W} = \frac{f}{Z_W} \tag{2.8}$$

The coordinates of '$P_W$' in the camera coordinate system can be calculated using the Euclidian transformation equation (Equation (2.5)). Hence the coordinates $(X_W, Y_W, Z_W)$ can be replaced with their corresponding camera coordinates denoted by $(X_C, Y_C, Z_C)$. Therefore, Equation (2.8), can be written as:

$$\lambda = \frac{u}{X_C} = \frac{v}{Y_C} = \frac{f}{Z_C} \tag{2.9}$$

From Equation (2.9), the projection of the 3-D target point to the image plane is expressed as:

$$(u,v) = \left( f\frac{X_C}{Z_C}, f\frac{Y_C}{Z_C} \right) \tag{2.10}$$

where $u = f\dfrac{X_C}{Z_C}$ and $v = f\dfrac{Y_C}{Z_C}$

The parameter which must be identified through camera calibration is the effective focal distance '$f$'.

Usually in computer vision literature, the origin of the image coordinate system is in the upper left hand corner of the image array (Heikkila and Silven, 1997). It is more convenient to express the image coordinates in terms of pixels. Therefore, the image coordinates $u$ and $v$ are multiplied by factors $D_u$ and $D_v$ (i.e. the pixel size) that specify the relationship between pixels and the physical object unit, millimeters (Heikkila, 2000). These factors can be typically obtained from the data sheets of the camera; however knowing their precise values is not necessary, because they are linearly dependent on the focal length and scale factor that are determined during calibration. Therefore, the corresponding image coordinates $(u', v')$ in pixels can be obtained from the projection $(u, v)$ by applying the following transformation:

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} D_u s u \\ D_v v \end{bmatrix} + \begin{bmatrix} U_0 \\ V_0 \end{bmatrix} \tag{2.11}$$

3. *Undistorted image coordinates to distorted image coordinates. (Nonlinear lens distortion model that accounts for radial and tangential distortions)*

The pinhole camera model is only an approximation of the real camera projection which provides a simple mathematical relationship between the object and image coordinates (Heikkila and Silven, 1997). In real cameras, perspective projection is insufficient for modeling the mapping between 3-D object and 2-D image coordinates precisely. Hence, it is not valid when high accuracy is required and, therefore, a more comprehensive camera model must be used. Ideally, the light rays coming from the scene should pass through the optical center linearly, but in practice, lens systems are composed of several

optical elements introducing nonlinear distortion to the optical paths and the resulting images (Heikkila, 2000). The most commonly used approach for correcting this lens distortion is to decompose the distortion into radial and tangential components. Usually, the pinhole model is a basis that is extended with some corrections for the systematically distorted image coordinates (Heikkila and Silven, 1997). In this study, a nonlinear fourth order radial distortion model and a second order tangential distortion model are used. The correction for the combined effect of both radial and tangential distortion can be expressed using the nonlinear lens distortion model as follows:

$a_c \approx a_d + F(a_d, \delta)$ where:

$$\mathrm{F}(a_d, \delta) = \begin{bmatrix} u_d F_{Radial} + F^x_{Tangential} \\ v_d F_{Radial} + F^y_{Tangential} \end{bmatrix} = \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} u_d\left(k_1 r_d^{\,2} + k_2 r_d^{\,4}\right) + \left(2p_1 u_d v_d + p_2\left(r_d^{\,2} + 2u_d^{\,2}\right)\right) \\ v_d\left(k_1 r_d^{\,2} + k_2 r_d^{\,4}\right) + \left(2p_2 u_d v_d + p_1\left(r_d^{\,2} + 2v_d^{\,2}\right)\right) \end{bmatrix}$$

$$r_d = \sqrt{u_d^{\,2} + v_d^{\,2}}$$

$$\delta = [k_1, k_2, p_1, p_2]^T$$

$$a_c = [u_c, v_c]$$

$$a_d = [u_d, v_d]$$

(2.12)

$\Delta u, \Delta v$ are the corrections for the combined effects of radial and tangential distortion. $a_c = [u_c, v_c]^T$ are the corrected image coordinates (corresponding to an ideal pinhole camera model) and $a_d = [u_d, v_d]^T$ are the distorted image coordinates.

Equation (2.12) represents an approximation for the corrected image coordinates from the distorted (or measured) image coordinates. Unfortunately there is no analytical closed-form solution for computing the distorted image coordinates, $a_d$, from the

corrected image coordinates, $a_c$. Several approximations for $a_d$ are proposed in the literature (e.g. Heikkila, 2000; Melen, 1994). In this study, $a_d$ is calculated by solving Equation (2.12) iteratively by applying Newton's method. The parameters which must be identified in this transformation are the radial distortion coefficients, $k_1$ and $k_2$ and the tangential distortion coefficients, $p_1$ and $p_2$.

4. *Distorted image coordinates to computer coordinate system (pixels)*

The distorted image coordinates $u_d$ and $v_d$ are measured relative to the projection centre of the lens. In computer vision applications, it is customary to specify the coordinates of pixels with respect to the upper left-hand corner of the image. Therefore, the transformation of the distorted image coordinates (measured in mm) to pixels in the computer reference frame is given by:

$$CX_I = \frac{u_d s}{P_x} + U_0, CY_I = \frac{v_d}{P_y} + V_0 \tag{2.13}$$

$(CX_I, CY_I)$- row and column numbers of the image pixel in computer frame memory

$(U_0, V_0)$- image center

$P_x$ - pixel size in X direction

$P_y$ - pixel size in the Y direction

Equation (2.13) can be expressed using the transformation matrix as follows:

$$\begin{bmatrix} CX_I \\ CY_I \end{bmatrix} = \begin{bmatrix} D_u s u_d \\ D_v v_d \end{bmatrix} + \begin{bmatrix} U_0 \\ V_0 \end{bmatrix} \tag{2.14}$$

where $D_u = \frac{1}{P_x}$ and $D_v = \frac{1}{P_y}$

In summary, a proper camera model for high accuracy computer vision applications can be derived by combining the pinhole camera model with the correction for radial and tangential distortion. The set of intrinsic parameters ($f$, $s$, $U_0$, $V_0$) is supplemented with the radial distortion coefficients, $k_1$ and $k_2$, and the tangential distortion coefficients, $p_1$ and $p_2$. These intrinsic parameters are also known as physical camera parameters since they have a certain physical meaning (Heikkila and Silven, 1997).

The mathematical transformation in GCC is represented as a flow chart in the next section, Section 2.5.2.

## 2.5.2 Flow Chart Representation

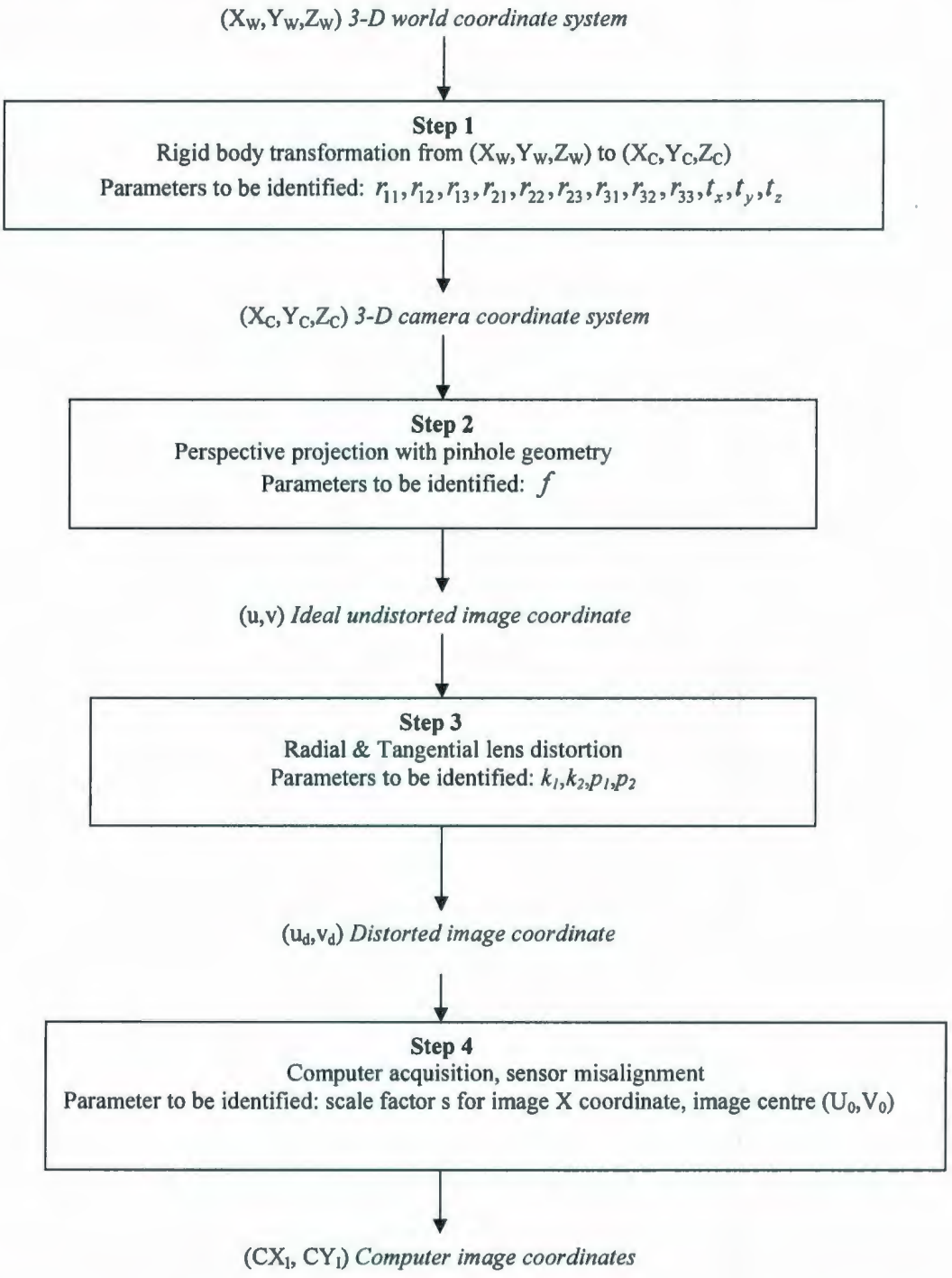



Figure 2.8: Four steps of transformation from 3-D world coordinate to computer image coordinate (Tsai, 1987)

Figure 2.8 represents the four steps of transformation from 3-D world coordinates to 2-D image coordinates. This flow chart can be explained briefly as follows: (1) Step 1 represents the rigid body transformation from the 3-D object coordinates in the world frame to 3-D object coordinates in the camera frame. The parameters that must be identified during this transformation are the nine rotational parameters and the three translational parameters, collectively referred to as extrinsic parameters. (2) Step 2 represents the transformation from the camera frame to the image plane giving ideal 2-D undistorted image coordinates. The parameter that is identified is the focal length. (3) Step 3 represents the transformation from undistorted image coordinates to distorted image coordinates as a result of radial and tangential distortion. The parameters that are identified are the four coefficients of radial and tangential distortion. (4) Step 4 represents the transformation from distorted coordinates in the image plane to the computer coordinate system which yields the coordinate locations in pixels. The parameters that must be identified are the scale factor and the image centre. The transformation from Step 2 to Step 4 identifies the camera's intrinsic parameters.

This procedure outlines the mathematical transformation used in this research for the purpose of camera calibration.

# Chapter 3

# Camera Calibration Techniques and Algorithms

## 3.1 Background and Motivation

This research work implements the techniques and algorithms reported in the paper by Heikkila, *Geometric Camera Calibration Using Circular Control Points* (Heikkila, 2000). This technique is considered to be representative of most techniques that are currently employed for camera calibration.

## 3.2 Discussion of Heikkila's Method

The technique of camera calibration presented by Heikkila in the paper *Geometric Camera Calibration Using Circular Control Points* involves a forward camera model

which converts the 3-D world coordinates of control points to distorted image coordinates and a backward camera model that transforms these distorted camera coordinates to lines-of-sight in the 3-D world coordinate system, or to the intersections of these lines with a known 2-D plane. The technique utilizes circular control points; however, the calibration procedure is also suitable for small points without geometry in which the radius is set to zero. Two types of lens distortions, namely radial and tangential distortion, are taken into consideration along with the pinhole camera model. The nonlinear lens distortion model used for correcting the distorted image coordinates is given by:

$$a_c = a_d + F_D(a_d, \delta),$$

where:

$$F_D(a_d,\delta) = \begin{bmatrix} \bar{u}_d\left(k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6 + \ldots\right) + \left(2p_1 \bar{u}_d \bar{v}_d + p_2\left(r_d^2 + 2\bar{u}_d^2\right)\right)\left(1 + p_3 r_d^2 + \ldots\right) \\ \bar{v}_d\left(k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6 + \ldots\right) + \left(p_1\left(r_d^2 + 2\bar{v}_d^2\right) + 2p_2 \bar{u}_d \bar{v}_d\right)\left(1 + p_3 r_d^2 + \ldots\right) \end{bmatrix} \tag{3.1}$$

$$a_d = [u_d, v_d]^T, a_c = [u_c, v_c]^T, \bar{u}_d = u_d - u_0, \bar{v}_d = v_d - v_0, r_d = \sqrt{\bar{u}_d^2 + \bar{v}_d^2}, \delta = [k_1, k_2, \ldots p_1, p_2, \ldots]^T$$

The distorted and corrected coordinates are denoted by $a_d$ and $a_c$, respectively. Since circular control points are used, a mathematical model for ellipse fitting is also presented in the procedure as the projection of a circle will be an ellipse during mapping. It is assumed that the camera model includes eight intrinsic parameters $\theta_{\text{int}} = [s, f, u_0, v_0, k_1, k_2, p_1, p_2]^T$ and six extrinsic parameters $\theta_{ext} = [t_x, t_y, t_z, \omega, \varphi, \kappa]^T$. In order to minimize the error between the observed and model coordinates, the distorted image coordinates are expressed in terms of their undistorted counterparts using an inverse distortion model. Since there is no analytic or closed form solution for the inverse

distortion problem, an approximate inverse model is found by taking the Taylor series of $F_D$ about $a_c$ which yields the following model:

$$a_d \equiv a_c - F_D^*(a_c, \delta)$$

where:

$$F_D^*(a_c, \delta) = \frac{1}{4k_1 r_c^2 + 6k_2 r_c^4 + 8p_1 \bar{v}_c + 8p_2 \bar{u}_c + 1} F_D(a_c, \delta), r_c = \sqrt{\bar{u}_c^2 + \bar{v}_c^2} \tag{3.2}$$

The parameters obtained using the inverse camera model are subsequently used to correct the distorted image coordinates. Due to the nonlinear nature of the camera model, an iterative search technique is used to produce an optimum solution. The calibration procedure suggested in (Heikkila, 2000) involves three steps and can be summarized as follows:

*Step 1*: *Initialization*: Before using an optimal estimator, initial estimates of the camera parameters are required to ensure that a global minimum can be achieved. In the initialization step, the aim is to produce the first, not the final, estimates of the camera parameters. It is often more reliable to use the nominal values for the focal length, scale factor and the image center as the initial estimates of the intrinsic parameters.

*Step 2*: *Iterative search*: In this step the parameters of the forward camera model are estimated by minimizing the weighted sum of the squared differences between the observations and the model. Numerical techniques such as the Lavenberg-Marquardt nonlinear optimization method are used to compute the optimal camera parameters.

*Step 3*: *Backward camera model*: This model involves the projection of the observed image coordinates back to 3-D coordinates. Once Steps 1 and 2 have been completed, the

parameters of the forward camera model are known and a set of distorted points $\{a_d(i)\}$ for arbitrary points $\{a_c(i)\}$, where $i$ = 1, ...., M can be produced. The inverse distortion model parameters are calculated and used for correcting the distorted image. Due to the approximations made in the derivation of the inverse distortion model, the parameter vector, $\delta$, must be adjusted for back projection using the following least-squares formulation:

$$\hat{\delta}' = \begin{bmatrix} B(1) \\ B(2) \\ \cdot \\ \cdot \\ B(M) \end{bmatrix}^{+} \begin{bmatrix} a_c(1) - a_d(1) \\ a_c(2) - a_d(2) \\ \cdot \\ \cdot \\ a_c(M) - a_d(M) \end{bmatrix} \tag{3.3}$$

where $\hat{\delta}'$ is the vector of the distortion parameters to be used in the backward camera model,

$$B(i) = \begin{bmatrix} \bar{u}_d(i) r_d^2(i) & \bar{u}_d(i) r_d^4(i) & \ldots & 2\bar{u}_d(i)\bar{v}_d(i) & r_d^2(i) + 2\bar{u}_d^2(i) & \ldots \\ \bar{v}_d(i) r_d^2(i) & \bar{v}_d(i) r_d^4(i) & \ldots & r_d^2(i) + 2\bar{v}_d^2(i) & 2\bar{u}_d(i)\bar{v}_d(i) & \ldots \end{bmatrix}, \tag{3.4}$$

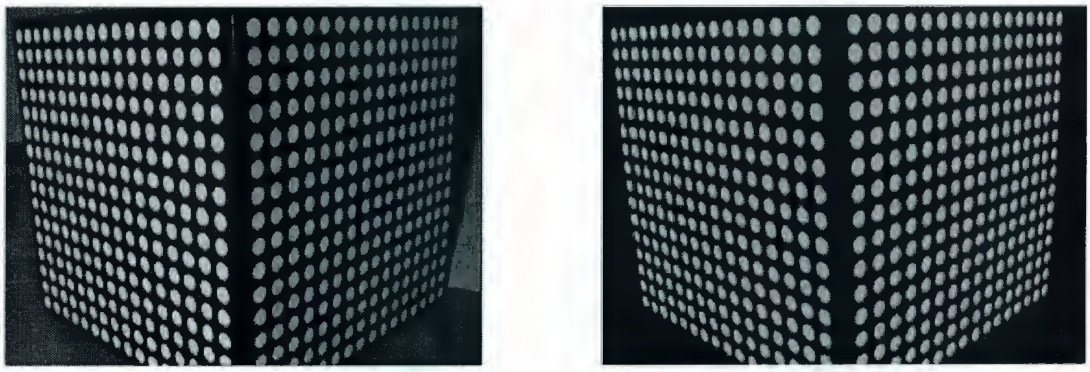
where $\bar{u}_d(i) = u_d(i) - u_0$, $\bar{v}_d(i) = v_d(i) - v_0$, and $r_d(i) = \sqrt{\bar{u}_d^2(i) + \bar{v}_d^2(i)}$.

### 3.2.1 Experiment Overview

Heikkila performed two experiments to test the proposed camera calibration procedure. First, the parameter estimates of the forward camera model were analyzed statistically using synthetic images and the results were compared with the outcome of a corresponding real image. Second, the precision of the inverse distortion model was evaluated by correcting and distorting random image coordinates.

*1. Calibrating the Forward Camera Model*

The tests were performed with 200 synthetic images and one real image that was used as a reference model. Figure 3.1 below shows the calibration images used for the experiments, one being the real image captured using an off-the-shelf monochrome CCD camera and the other being the corresponding synthetic image.



(a) (b)

Figure 3.1: Calibration images (a) Real image (b) Synthetic image (Heikkila, 2000)

The calibration object consists of two perpendicular planes each with 256 circular control points. The centers of these control points were located using the moment and curvature preserving ellipse detection technique and renormalization conic fitting. The calibration procedure was first applied to estimate the camera parameters based on the real image. The synthetic images were then produced using ray tracing with the camera parameters obtained from calibration and the known 3-D model of the control points. In order to make the synthetic images better correspond to the real images, their intensity values were perturbed with additive Gaussian noise $(\sigma = 2)$, and blurred using a 3 by 3 Gaussian filter $(\sigma = 1 pixel)$ as shown in Figure 3.1(b). Three different calibration methods were applied separately for all control point sets. The first method is the traditional camera calibration approach which does not assume any geometry for the control points. In the second method, circular geometry is utilized and all the observations are equally weighted. In the third method, each observation is weighted by the inverse of the observation error covariance matrix $C_e$. The estimated bias and standard deviation of the intrinsic parameters for all three methods were reported. Confidence intervals with a 95 percent confidence level were also presented. It was noted that control points with circular geometry can reduce the bias significantly.

*2. Calibrating the Reverse Distortion Model*

In this experiment, the accuracy of the reverse distortion model was tested with a wide range of distortion parameters. For each $\hat{\delta}$, the third step of the calibration procedure was applied to solve for the parameter vector $\hat{\delta}'$. Then 10,000 random points were

generated inside the image region which were corrected and distorted using the distortion models mentioned in equations (3.1) and (3.2). The difference between the resulting coordinates and the original coordinates was examined with respect to the different values of the distortion parameters. The curves showing the root mean square error in the image coordinates as a function of distortion parameters $k_1$ and $k_2$ is also reported in the paper.

The paper also discusses the various error sources that reduce the theoretical accuracy. Depending on the study done, the paper concludes that an accuracy of 1/50 of the pixel size is achievable with the presented technique if error sources, such as line jitter and chromatic aberration are eliminated. The tests with synthetic images indicate improvements in the calibration results in limited error conditions. In real images, the suppression of external error sources becomes a prerequisite for successful calibration.

# Chapter 4

# Implementation and Assessment of Accuracy

In the field of machine vision, Geometric Camera Calibration refers to the experimental determination of a set of parameters which describe the image formation process for a given analytical model of the machine vision system. This study implements the camera calibration techniques and algorithms described in (Heikkila, 2000). The first problem that is addressed in this thesis is to find out whether the algorithm reported in (Heikkila, 2000), works properly in estimating the camera parameters. As a second step, an accuracy assessment is performed on Heikkila's method. Accurate camera calibration is necessary in many vision based applications that involve quantitative measurements. In general, no matter the calibration method used, the accuracy of the calibration depends on the known geometry of the targets which the lens or camera views. The camera support

during calibration should be rigid and shielded from vibrations; thermal variables should be negligible (Slama, 1980). Accuracy is the degree of closeness of a measured or calculated quantity to its actual (true) value. In this study, the accuracy analysis is performed using a synthetic model and deals with the fundamental question, *"What is the accuracy required on the target points for optimal camera calibration?"* In order to generate a synthetic model, the assumption is that the 3-D coordinates of a certain number of target points are known with respect to the world coordinate frame. In GCC, these 3-D target points are mapped to 2-D image points by applying the forward camera model. The unknown camera parameters are then solved by minimizing the discrepancy between the measured 2-D image points and those obtained from the model.

The control points of the calibration target can be either coplanar (e.g. a printed checkerboard pattern as shown in Figure 1.3) or three dimensional (e.g. two orthogonal planes as depicted in Figure 1.2). In the case of a coplanar target, multiple images captured from different positions and orientations are required. If the control point structure is three dimensional, the advantage is that only one image is required for calibration. In order to generate a synthetic image of a calibration target, the configuration of the imaging system must first be specified.

# 4.1 Implementation and Validation of Heikkila's Method

## 4.1.1 Problem Statement

The first problem that is addressed in this thesis is to determine whether the algorithm reported in Heikkila (2000), works properly in estimating the camera parameters assuming ideal data (i.e. no measurement errors). This involves generating a synthetic or virtual image by defining target points in the world coordinate system and then mapping these 3-D object coordinates to 2-D image coordinates using a set of known extrinsic and intrinsic camera parameters. The specific parameters used in this study correspond to a commercially available camera, Lumenera (Lm135 1.4 megapixel industrial USB 2.0 camera, Figure 4.1) and lens assembly.


Figure 4.1: Camera system, Lumenera

#### 4.1.1.1 Calibration model

The validation procedure involves generating a synthetic camera calibration model that includes:

1. Target points in the world coordinate system:
    - An array of 1200 target points in three parallel planes (20x20x3) with a 5.715mm spacing between each target point in the X direction, 3.810mm in the Y direction and 6.35mm in the Z direction (Figure 4.2). This spacing was selected to cover the entire field-of-view of the camera.

2. Position and orientation of target points relative to the camera coordinate system:
    - Rotation of target about Z axis, $\alpha_r = 0$
    - Rotation of target about Y axis, $\beta_r = 0$
    - Rotation of target about X axis, $\gamma_r = 0$
    - X position of target relative to camera coordinate system, (trans_x) = -60mm
    - Y position of target relative to camera coordinate system, (trans_y) = -40mm
    - Z position of target relative to camera coordinate system, (trans_z) = 800mm

3. Effective focal length of the lens:
    - $f$ = 35mm

4. Image center in pixels measured relative to the image coordinate system with origin in the upper left-hand corner of the image. The pixel dimensions in the X and Y direction is 0.00465mm as provided by the camera manufacturer:

   - X image center in pixels, $U_0$ = 696
   - Y image center in pixels, $V_0$ = 520

5. Radial and Tangential distortion coefficients:

   - Radial distortion coefficients, Rad1 ($k_1$) = 0.0001, Rad2 ($k_2$) = 0.00005
   - Tangential distortion coefficients, Tan1 ($p_1$) = 0.0001, Tan2 ($p_2$) = 0.00005

Figure 4.2 below is a diagrammatic representation of the synthetic camera calibration model.

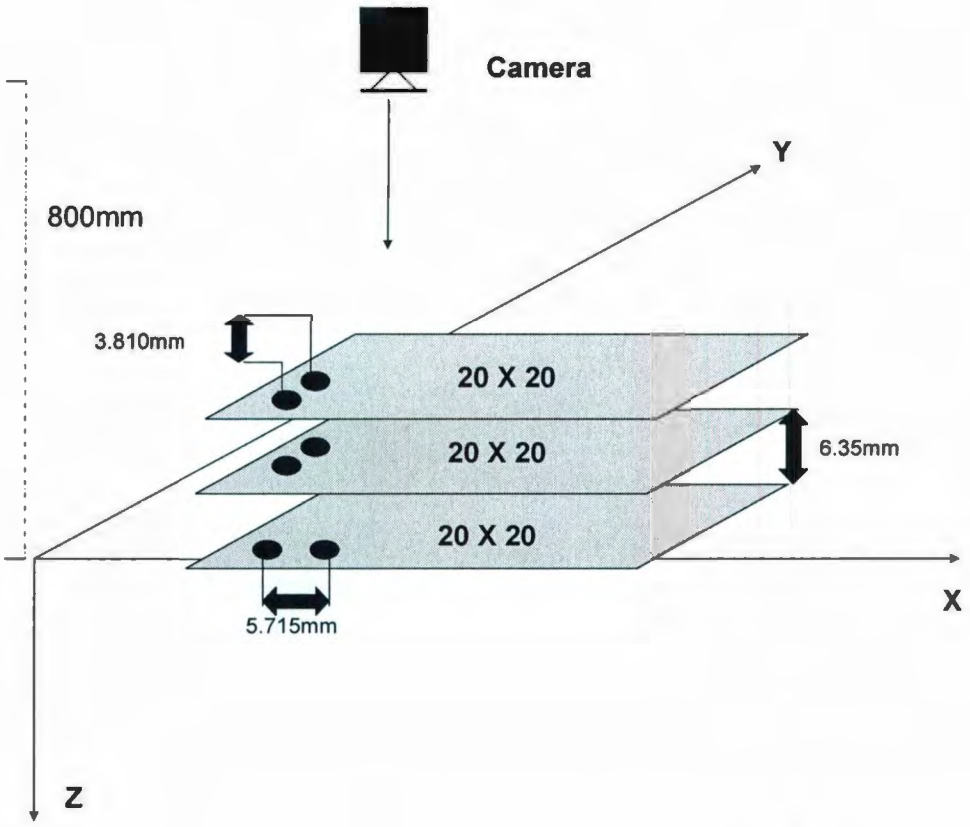



Figure 4.2: Diagrammatic representation of the synthetic camera calibration model

### 4.1.1.2 Algorithm

The algorithm used to validate the calibration procedure can be summarized as follows:

Step 1: Define the world coordinates of target points.

Step 2: Define the rotation matrix relating the world coordinate system to the camera coordinate system. Rotate and translate the target points to obtain the points relative to the camera coordinate system.

Step 3: Assuming a pinhole camera model, calculate the coordinates of the undistorted image coordinates of the target points (measured in mm) relative to the image center of the camera.

Step 4: Calculate the radial and tangential distortion using the nonlinear lens distortion model and compute the distorted image coordinates of the target points.

Step 5: Introduce the scale factor and convert the units of the distorted image coordinates from mm to pixels using the pixel dimensions provided by the camera manufacturer.

Step 6: Add the image center offset to the distorted image coordinates to obtain the image coordinates relative to the computer coordinate system measured relative to the upper left hand corner of the image.

Step 7: Using the distorted computer image coordinates obtained in Step 6 and the world coordinates of the target points defined in Step 1, perform the camera calibration by applying Heikkila's algorithm.

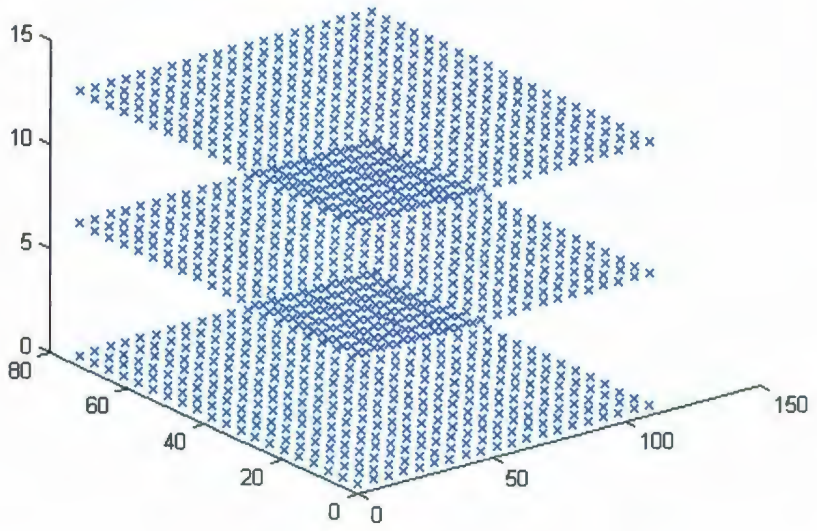
Step 8: Calculate the inverse model parameters and determine the corrected image coordinates.

Step 9: Compare the original undistorted image coordinates of the target points in pixels with the corrected image coordinates and plot the histograms of the differences.

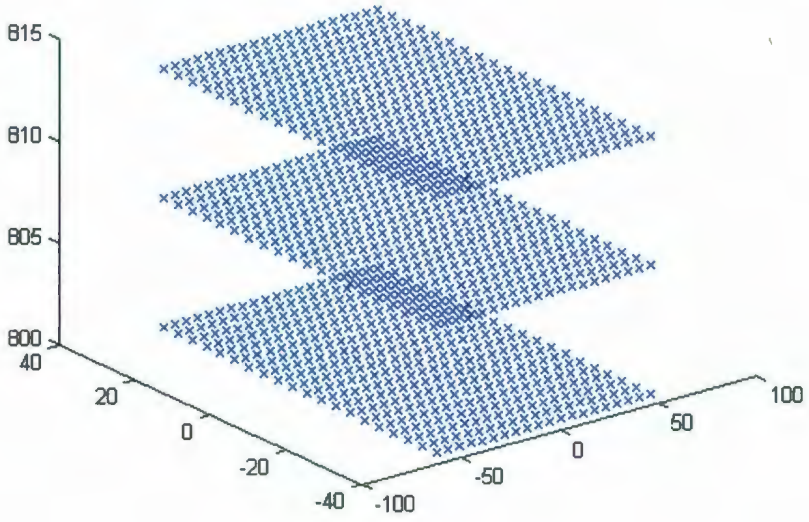
Step 10: Finally, compare the intrinsic and extrinsic parameters used to generate the synthetic image (Section 4.1.1.1) with those provided by Heikkila's calibration algorithm in Step 7 above.

### 4.1.2 Results and Discussions

Figures below represent the calibration target, synthetic images and the histogram plot of error in pixels in the X and Y directions obtained by following the steps described in the previous section. Note that the lens distortion is relatively weak as can be observed in Figure 4.4. The distortion parameters used are typical of a commercially available industrial camera with a relatively long focal length.

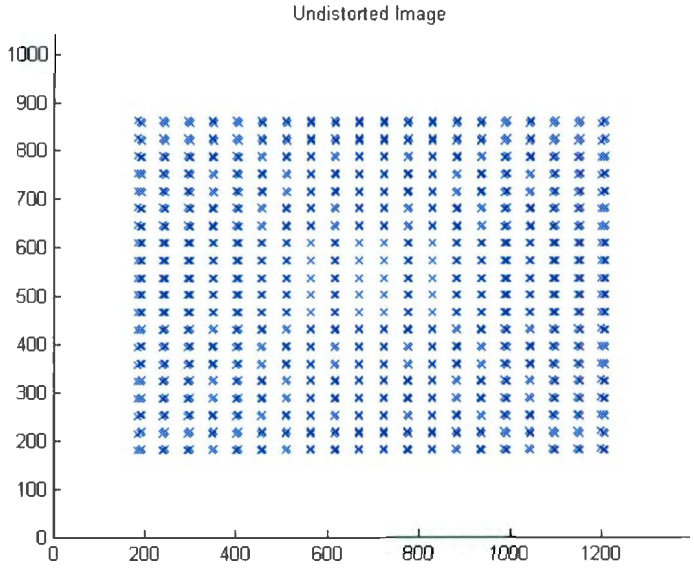



(a)

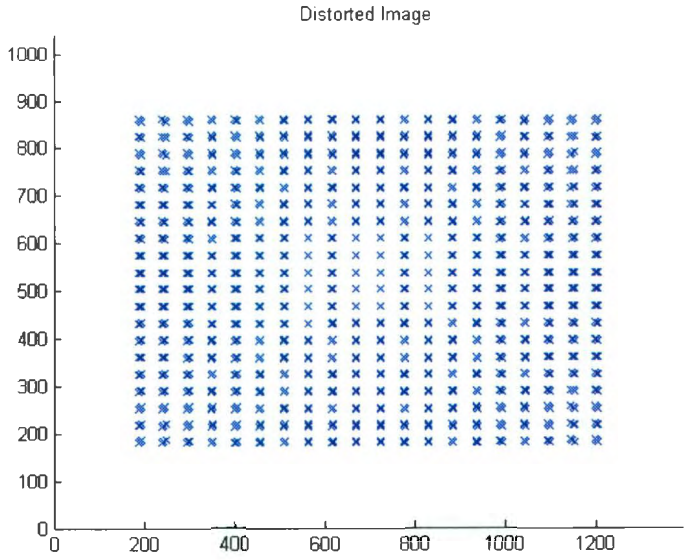



(b)

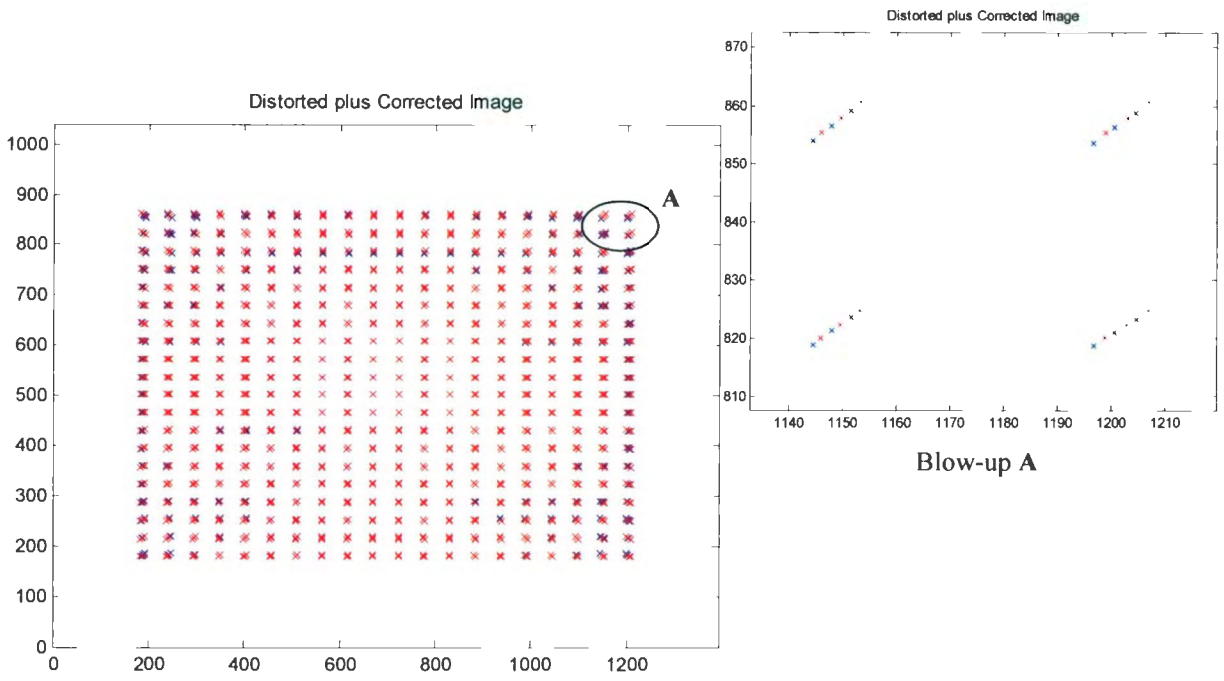
Figure 4.3: Calibration targets (in mm)
(a) Defined target points (b) Target points after undergoing rotation and translation


Undistorted Image
1000
900
800
700
600
500
400
300
200
100
0
0
200
400
600
800
1000
1200

(a)


Distorted Image
1000
900
800
700
600
500
400
300
200
100
0
0
200
400
600
800
1000
1200

(b)




(c)

Figure 4.4: Synthetic images (in pixels)
(a) Undistorted image (b) Distorted image after adding radial and tangential distortion (c) Distorted plus corrected image; blue: distorted, red: corrected (correction done using inverse model parameters)

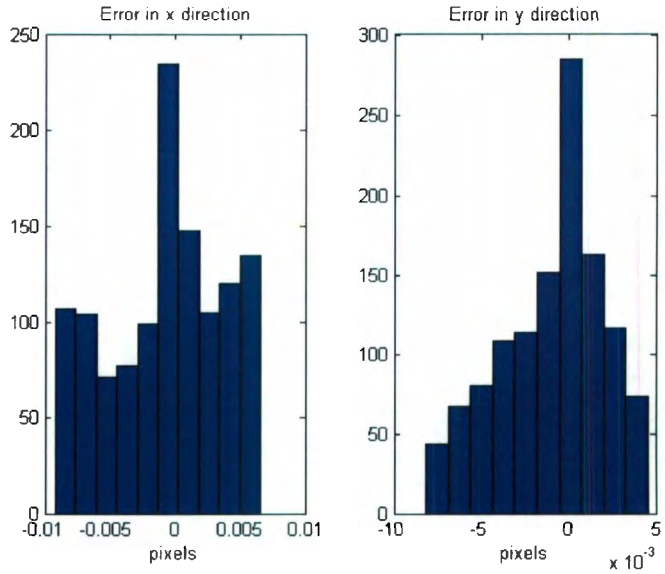



Figure 4.5: Error in pixels; histogram plots representing the difference between the original and the corrected image coordinates (error in X and Y direction)

Table 4.1 given below presents the calibration results of the model explained above.

**Table 4.1: Validation results**

| Parameters | Symbol | Theoretical Value | Calibration Results |
|---|---|---|---|
| Scale factor | $s$ | 1.0 | 1.0000 |
| Effective focal length | $f$ (mm) | 35.0 | 34.9999 |
| Principal point | $U_0$ (pixels) | 696 | 695.8728 |
| | $V_0$ (pixels) | 520 | 519.7194 |
| Radial distortion | $k_1$ ($mm^{-2}$) | 1.0e-4 | 9.788726e-005 |
| | $k_2$ ($mm^{-4}$) | 5.0e-5 | 5.046853e-005 |
| Tangential distortion | $p_1$ ($mm^{-1}$) | 1.0e-4 | 1.000345e-004 |
| | $p_2$ ($mm^{-1}$) | 5.0e-5 | 5.002318e-005 |
| Alpha | $\alpha_r$ (degrees) | 0 | 5.3713e-006 |
| Beta | $\beta_r$ (degrees) | 0 | 9.6931e-004 |
| Gamma | $\gamma_r$ (degrees) | 0 | -0.0021 |
| trans_x | $t_x$ (mm) | -60.0 | -59.9865 |
| trans_y | $t_y$ (mm) | -40.0 | -39.9702 |
| trans_z | $t_z$ (mm) | 800.0 | 800.0024 |

The results in Table 4.1 indicate that the calibration results for each parameter closely match the theoretical values used to generate the synthetic image. This implies that from a theoretical point of view the calibration algorithm accurately estimates the intrinsic and extrinsic camera parameters.

## 4.2 Accuracy Assessment

### 4.2.1. Problem Statement

The accuracy of the data acquisition process plays an important role when evaluating the performance of the camera calibration algorithm (Guendouz et al., 2006). The precision of the estimated camera parameters depends on the positional accuracy of the 3-D target points as well as on the technique employed to identify the 2-D image points. Hence, it can be inferred that uncertainty in the 3-D coordinates of the target points can potentially yield inaccurate results during camera calibration.

In an effort to better understand the influence of measurement error on camera calibration, this section of the thesis focuses on determining the required accuracy on the target points for optimal camera calibration. This is accomplished by adding uncertainty on the target point locations through randomly generated noise and studying the influence of the measurement error on the estimated camera parameters.

### 4.2.2 Accuracy Assessment Procedure

The accuracy assessment procedure involves adding an uncertainty of 0.1mm on each of the target points through randomly generated noise using the same synthetic camera calibration model depicted in Figure 4.6. Previously it was assumed that the exact coordinates of the 3-D target points were known relative to the world coordinate system.

In this section an error bound or tolerance is introduced on the position of the target points. It is assumed that each target point has an equal probability of being located somewhere within an error bound of +/-0.05 mm in each of the three orthogonal directions.

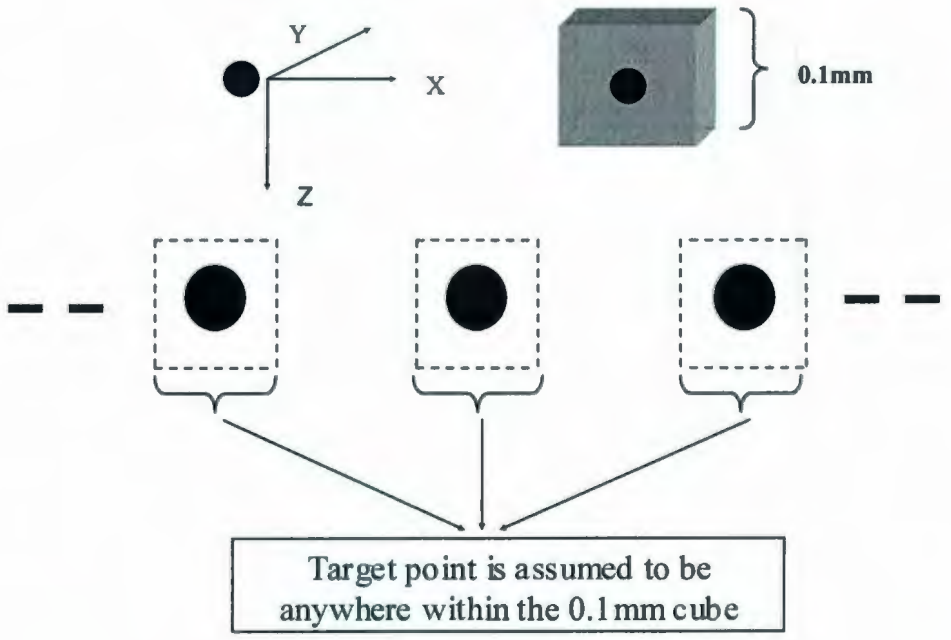



Figure 4.6: Diagrammatic representation showing addition of uncertainty

### 4.2.2.1 Algorithm

In order to verify how much accuracy is required on the target locations, the following study is performed:

Step 1: Add randomly generated noise on each of the target point coordinates as defined in the world coordinate system. In this manner each target point has an equal probability of falling anywhere within a 0.1 mm "uncertainty" cube.

Step 2-6: Calculate the distorted image coordinates of the target points in pixels relative to upper left hand corner of image (i.e. relative to the computer coordinate system) as described in Section 4.1.1.2.

Step 7: Perform the calibration by applying Heikkila's calibration algorithm.

Step 9: Repeat Step 7 and compute the mean and variance of the intrinsic and extrinsic parameters provided by the calibration algorithm. Continue until a constant value of variance is observed for each of the parameters (in this case 400 iterations are required).

Step 10: Plot the histogram of each of the parameters.

### 4.2.3 Results and Discussions

Table 4.2 summarizes the results of the study presented in the previous section. The various performance indices are defined as follows:

- The *Variance* and *Mean* for each parameter was computed by following the steps outlined above with an uncertainty of +/-0.05mm added onto each of the target points.
- The *Standard Deviation* is the square root of the *Variance*.

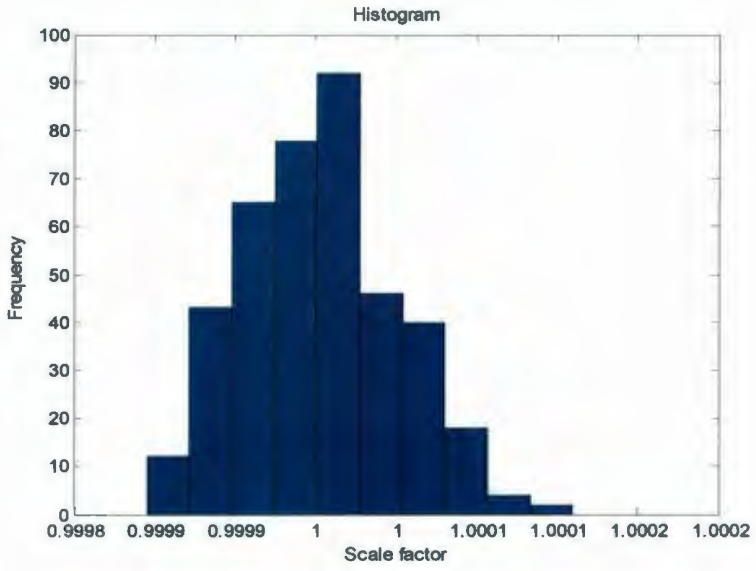
- The *Absolute Accuracy* is the 95% confidence limit which corresponds to twice the *Standard Deviation.*
- The *Relative Accuracy* is the *Absolute Accuracy* expressed as a percent of the *Mean* value. Note that *Relative Accuracy* is not defined for parameters that have a nominal value of zero.

**Table 4.2: Calibration results with an error bound of 0.1mm**

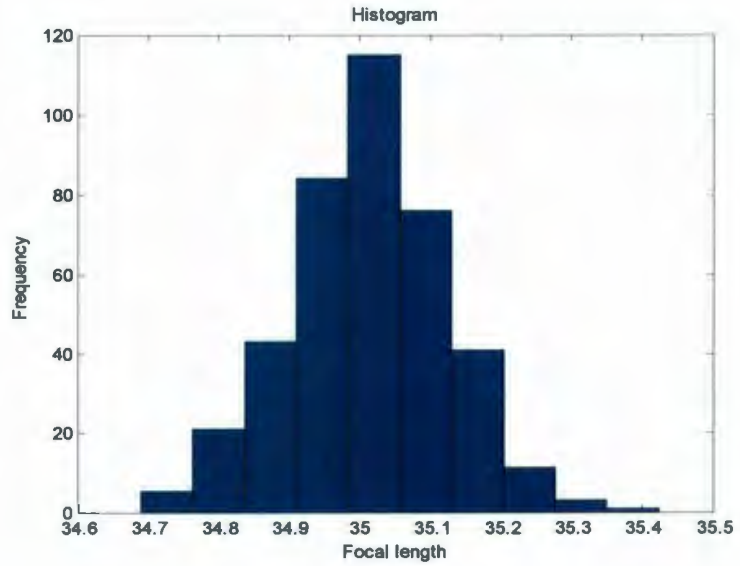
| Parameters | Variance | Standard Deviation | Mean | Absolute Accuracy | Relative accuracy |
|---|---|---|---|---|---|
| s | 2.26E-09 | 4.75E-05 | 1.0000 | 9.50E-05 | 0.0095% |
| $f$(mm) | 0.0125 | 0.112 | 35.0105 | 2.24E-01 | 0.6398% |
| $U_0$ (pixels) | 267.3676 | 16.3514 | 696.2667 | 3.27E+01 | 4.6969% |
| $V_0$ (pixels) | 132.4768 | 11.5099 | 518.9618 | 2.30E+01 | 4.4357% |
| $k_1(mm^{-2})$ | 3.01E-09 | 5.48E-05 | -2.91E-07 | 1.10E-04 | - |
| $k_2(mm^{-4})$ | 3.74E-11 | 6.11E-06 | 1.99E-08 | 1.22E-05 | - |
| $p_1(mm^{-1})$ | 3.11E-10 | 1.76E-05 | 1.62E-06 | 3.53E-05 | - |
| $p_2(mm^{-1})$ | 5.23E-10 | 2.29E-05 | -3.57E-07 | 4.58E-05 | - |
| Alpha, $\alpha_r$ | 1.66E-06 | 0.0013 | 9.23E-05 | 2.60E-03 | - |
| Beta, $\beta_r$ | 0.0153 | 0.1235 | -0.0027 | 2.47E-01 | - |
| Gamma, $\gamma_r$ | 0.0077 | 0.0877 | -0.0079 | 1.75E-01 | - |
| trans_x, $t_x$ | 3.0231 | 1.7387 | -60.028 | 3.48E+00 | -5.7929% |
| trans_y, $t_y$ | 1.4984 | 1.2241 | -39.8896 | 2.45E+00 | -6.1374% |
| trans_z, $t_z$ | 6.7001 | 2.5884 | 800.24 | 5.18E+00 | 0.6469% |

It is observed that if the locations of the 3-D control points are not precise, the estimates of the camera parameters are inaccurate. The camera parameters appear to

follow a Gaussian distribution centered about the mean or nominal value as shown in the histogram plots below.

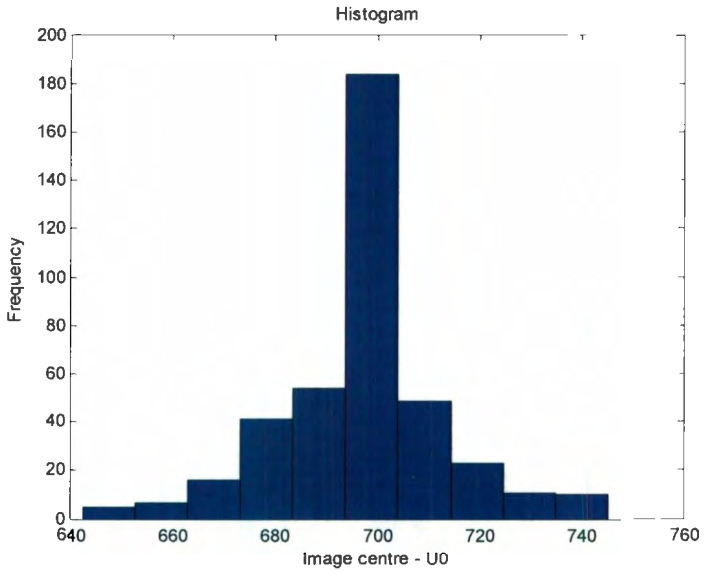



(a)

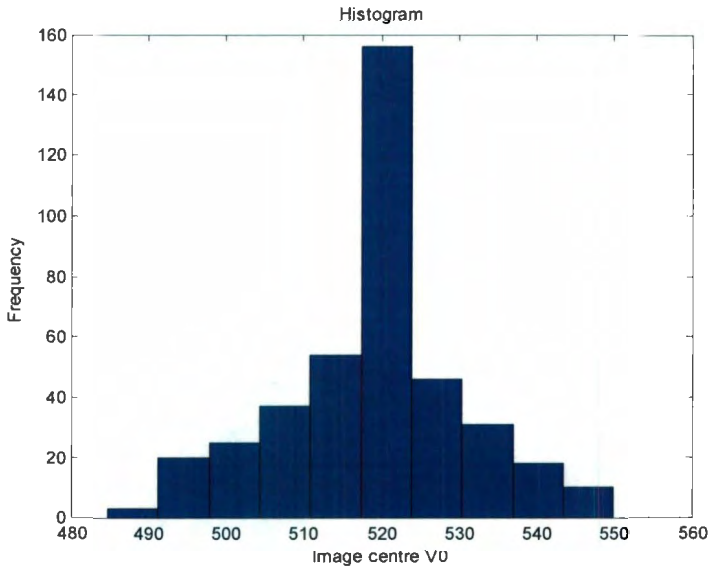



(b)

Figure 4.7: Variation in the scale factor and focal length observed for an uncertainty level of +/-0.05 mm (a) Scale factor (b) Focal length

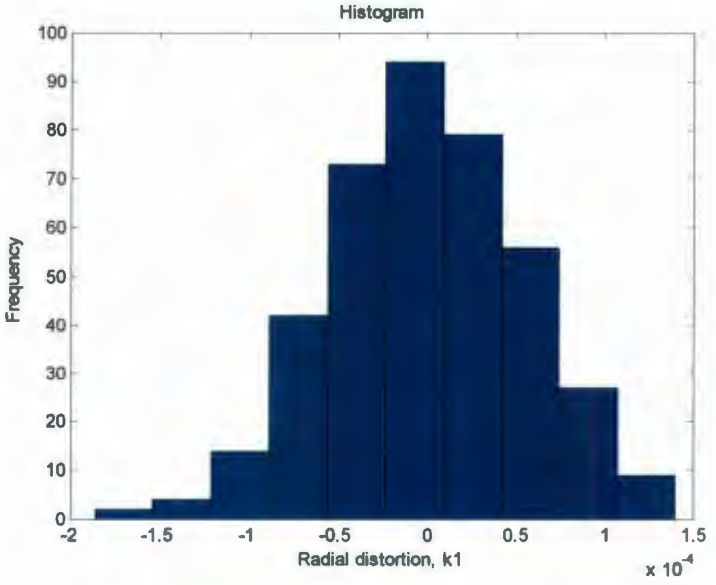



(a)

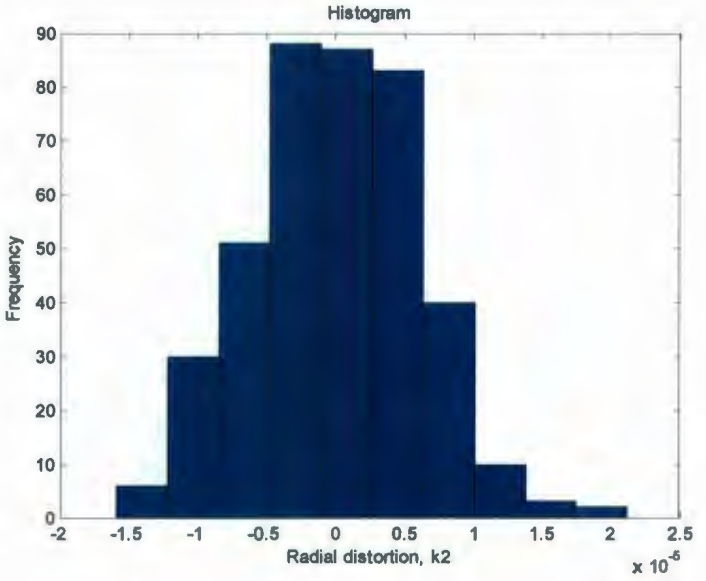



(b)

Figure 4.8: Variation in the image centre observed for an uncertainty level of +/-0.05 mm (a) Image centre, $U_0$ (b) Image centre, $V_0$

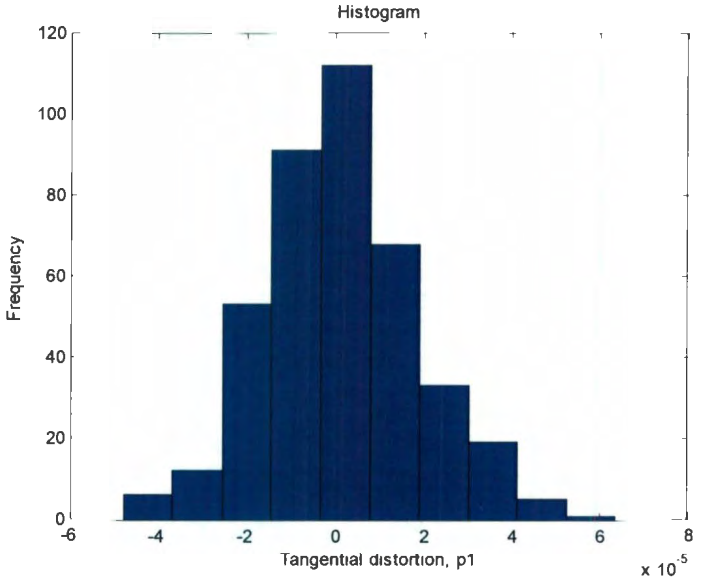



(a)

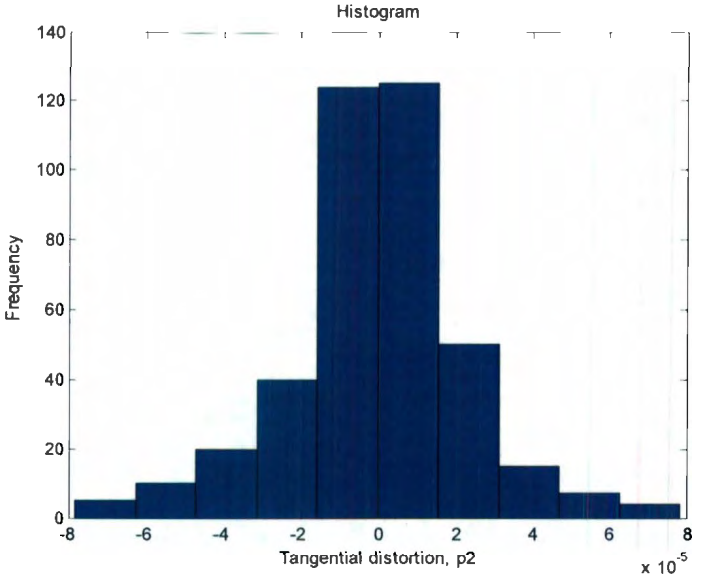



(b)

Figure 4.9: Variation in the radial distortion observed for an uncertainty level of +/-0.05 mm (a) Radial distortion, $k_1$ (b) Radial distortion, $k_2$

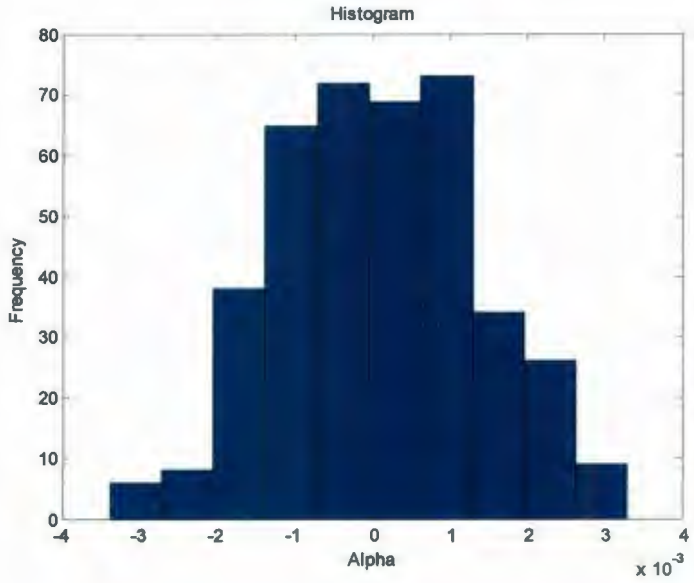



(a)

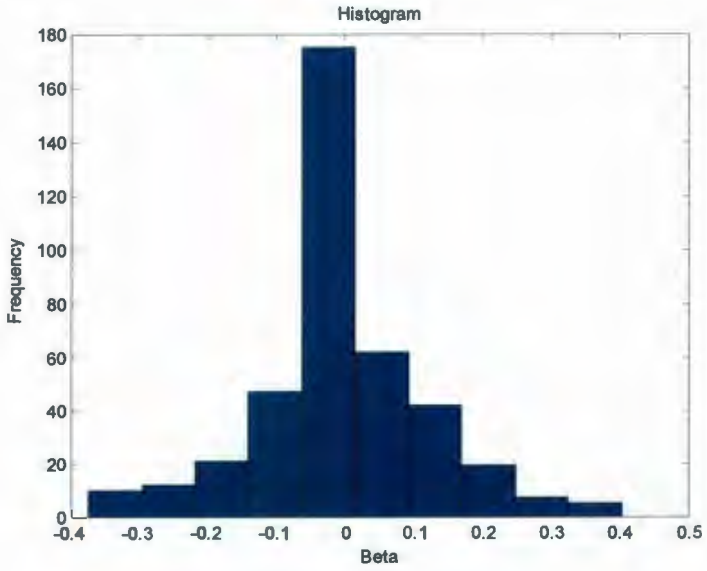



(b)

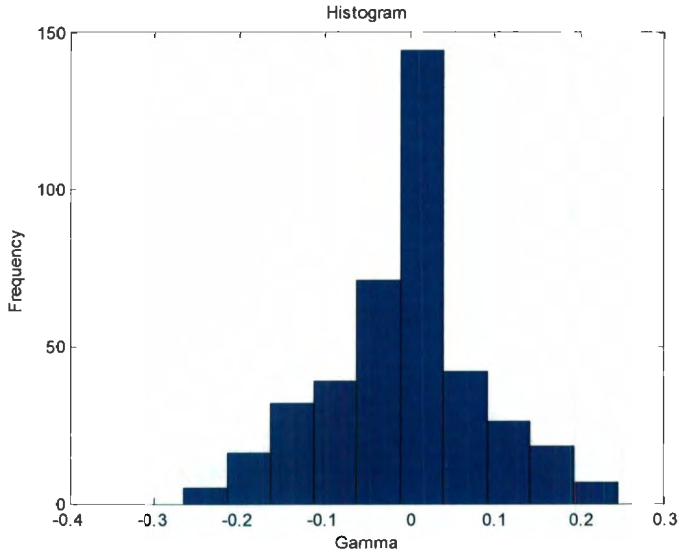
Figure 4.10: Variation in the tangential distortion observed for an uncertainty level of +/- 0.05 mm (a) Tangential distortion, $p_1$ (b) Tangential distortion, $p_2$




(a)

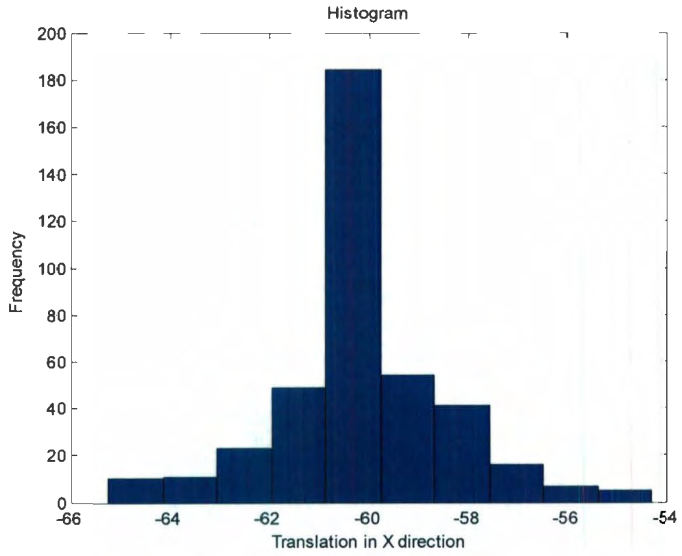



(b)

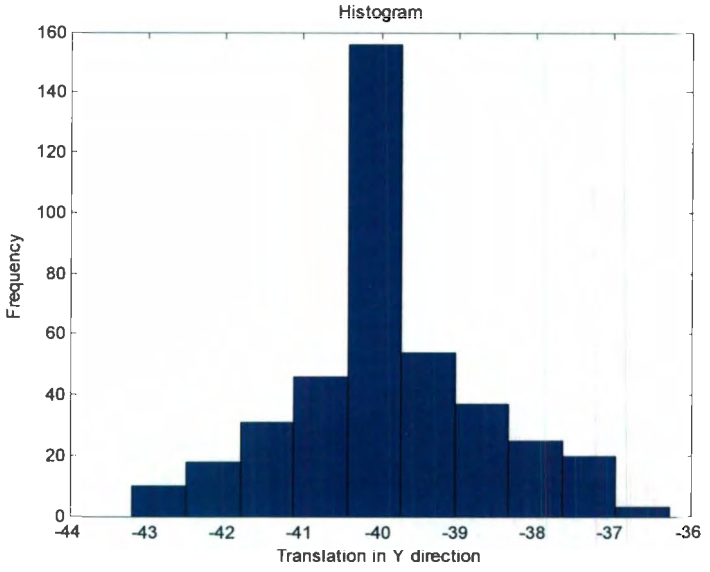



(c)

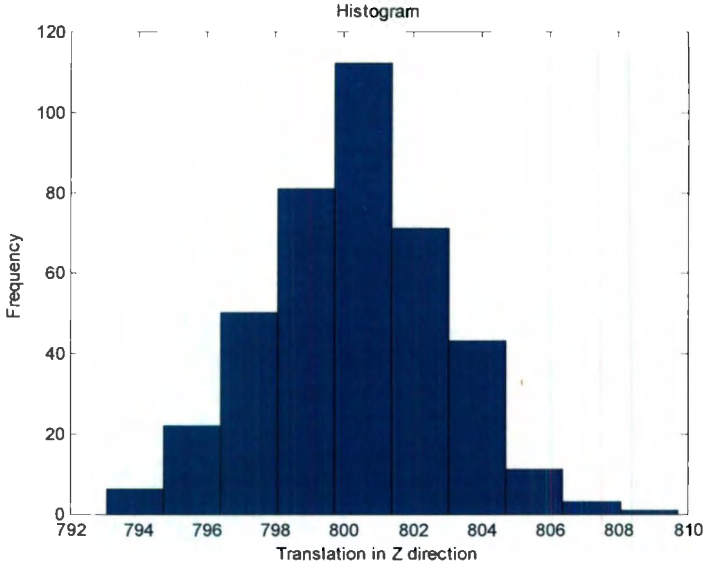
Figure 4.11: Variation in the rotational parameters observed for an uncertainty level of +/-0.05 mm (a)Alpha (b) Beta (c) Gamma




(a)




(b)




(c)

Figure 4.12: Variation in the translational parameters observed for an uncertainty level of +/-0.05 mm (a) trans_x (b) trans_y (c) trans_z

The figures shown above represent the histogram plots of the camera calibration parameters showing the mean and the standard deviation which characterizes the dispersion of data about the mean value when a random error is added onto the target points. A low standard deviation indicates that the data is clustered around the mean value, whereas a high standard deviation indicates that the data is widely spread with significant deviation from the mean. Note that the mean value corresponds to the predefined or nominal values of the parameters taken into consideration in the calibration model. It can be observed that the parameter values follow a Gaussian or Normal distribution. For example, Figure 4.7(b) shows that the values of focal length obtained with an uncertainty of +/-0.05 mm added to target points closely follows a Gaussian distribution with a mean value of 35mm. Upon close inspection, it can be observed that the statistical dispersion of distortion parameters is high in comparison to the other intrinsic parameters. It should be noted that for the distortion parameters, the order of magnitude of the variation is close to the actual value of the parameter. An important conclusion can be drawn from this; namely, that a very high accuracy is required on the target points in order to accurately determine the distortion parameters. With respect to the extrinsic parameters, the histogram plot of the first rotational parameter (i.e.; "$\alpha_r$" the rotation about the Z axis) is relatively accurate when compared to the other two rotational parameters. This is due to the fact that a relatively small change in "$\alpha_r$" results in a relatively large change in the observed image.

To conclude, the dimensional accuracy of the calibration target plays an important role in geometric camera calibration. The calibration results and accuracy analysis

presented in this chapter prove that the camera calibration techniques and algorithms described by Heikkila work properly in theory, but not necessarily in practice. In a real world machine vision application it is extremely difficult to obtain accurate or repeatable results even though current camera calibration techniques and algorithms are sound from a theoretical point of view.

# Chapter 5

# Analysis Using Design of Experiments

This chapter focuses on the task of quantitatively analyzing the problems associated with geometric camera calibration using the systematic Design of Experiments (DOE) approach. It has already been mentioned and demonstrated in Chapter 4 that it is extremely difficult to obtain accurate and repeatable results in practice, although, from a theoretical point of view, the camera calibration techniques and algorithms currently in use are sound. As an outcome of the study performed in this chapter, significant findings are reported including large interactions between certain factors that help explain the difficulties in obtaining accurate calibration results in many real-world machine vision applications where high accuracy is essential.

## 5.1 Design of Experiments - An Overview

Design of Experiments (DOE) is a structured, organized, and theoretically sound statistical approach for determining the relationship between factors (parameters) affecting a process and the output of that process (response). This method was first developed in the 1920s and 1930s, by Sir Ronald A. Fisher, the renowned mathematician and geneticist.

In simplest terms, DOE is a methodology for systematically applying statistics to experimentation. The DOE approach involves designing a series of structured tests in which planned changes are made to the input variables of a process or system and the effects of these changes on a pre-defined output are then assessed. Design of Experiments lets experimenters develop a mathematical model that predicts how input variables interact to create output variables or responses in a process or system. Design of Experiments methods are also very useful as a strategy for building mechanistic models and they have the additional advantage that no complicated calculations are needed to analyze the data produced from the designed experiment (Lye, 2003). The theoretical study and practical usage / applications of DOE are described thoroughly in (Cochran and Cox, 1992; Hicks and Turner, 1999; Mason et al., 2003; Montgomery, 2008; Myers and Montgomery, 2009).

Design of Experiments can be used for a wide range of experiments for various purposes including nearly all fields of engineering (e.g. chemical, manufacturing, civil, environmental, etc.), science (e.g. biology, agriculture, etc.) and to a lesser extent in

business marketing studies (Lye, 2005). It is, however, rarely used in fields such as intelligent systems; a multidisciplinary area of study involving computer vision, machine vision, robotics, artificial intelligence, computing / neural networks, etc.

### 5.1.1 When to use DOE?

Design of Experiments can be used to find answers in various situations such as "what is the main contributing factor to a problem?"; "how well does the system/process perform in the presence of noise?"; "what is the best configuration of factor values to minimize variation in a response?"; (The Quality Portal, [Online]).

In general, by using DOE, one can (Lye, [Online]):

- Learn about the process involved in the investigation
- Screen important variables
- Build a mathematical model
- Obtain prediction equations
- Optimize the response (if required)

The DOE approach is more efficient than the one factor at a time approach, because it requires fewer test runs and it plans for all possible dependencies or interactions between the levels or range of factors taken under consideration. The factorial approach of DOE is the most modern, theoretically sound and efficient approach. It is applied extensively in industrial research and development and process improvement.

The basic techniques of DOE consist of a combination of experimental design, analysis of variance (ANOVA), and regression analysis. An experiment is a test or a series of tests in which purposeful changes are made to the *input variables (design variables) or factors* of a system so that the reasons for changes in the *output response(s) or response or dependent variables* are observed or identified (Lye, [Online]). An experimental design is a strategy to gather empirical knowledge; i.e. knowledge based on the analysis of experimental data and not on theoretical models. It can be applied whenever the goal is to investigate a phenomenon in order to gain understanding or improve performance (CAMO, [Online]). Building a design means carefully choosing a small number of experiments that are to be performed under controlled conditions. There are four interrelated steps in building a design:

1. Define an objective to the investigation; e.g. better understand or sort out important variables or find optimal values.
2. Define the variables that will be controlled during the experiment ("factors or design variables") and their levels or ranges of variation.
3. Define the variables that will be measured to describe the outcome of the experimental runs ("response variables or dependent variables") and examine their precision.
4. Among the available standard designs, choose the one that is most compatible with the objective, the number of input factors, and the precision of the measurements. The design selected should also have a reasonable time and budget especially in practical experiments.

The statistical significance of the experimental design is tested using Analysis of Variance (ANOVA). The analysis of variance is a mathematical process for separating or partitioning the total variability of a group of observations into its assignable causes and setting up various significance tests (NIST/SEMATECH, [Online]). The goal of regression analysis is to determine the values of parameters for a function that cause the function to best fit the set of data observations provided (NLREG, [Online]). In DOE, it is used to obtain the prediction model.

## 5.2 Problem Definition

A statistical factorial design of experiments is an efficient means to simultaneously study the effect of several input factors on an output response and determine the optimum settings for them. In addition, any potential interactions between the factors can be evaluated (NIST/SEMATECH, [Online]). The problem with geometric camera calibration based on the accuracy analysis study presented in Chapter 4, is analyzed here by applying the concepts of design of experiments. The procedure for applying DOE to better understand the question of accuracy with geometric camera calibration can be summarized as follows:

1. Set the objective and process variables, i.e. input factors (choose factor levels) and response variables.
2. Choose a suitable experimental design depending on the objective of the experiment and the number of factors to be investigated.

3. Execute the chosen experimental design and analyze the experiment by selecting model factors, model reduction, finding significant factors through ANOVA analysis, forming model equations and analyzing residuals for model adequacy checking.
4. Interpret the results obtained from the model graph plots.
5. Validate the model plots by comparing with the results obtained using the MATLAB calibration model under study.

A more detailed description on the methodology undertaken for the study, the discussion on the results and the validation approach is described in Section 5.2.1, Section 5.3 and Section 5.4.

## 5.2.1 Experimental Design

This section of the chapter discusses the use of DOE to investigate the influence of seven intrinsic parameters on the accuracy of the camera calibration procedure. The calibration model based on a 3-D calibration target consisting of three planes of control points as explained in Chapter 4 was considered for the statistical experiment. Seven intrinsic parameters namely; focal length, image centre ($U_0$,$V_0$), radial ($k_1$, $k_2$) and two tangential distortion ($p_1$, $p_2$) coefficients were chosen as factors. The response was taken as the sum of the square of the error in the image coordinates caused by different combinations and levels of the factors considered. The error was calculated as the change in the image coordinate locations relative to the actual image coordinates due to the change in the

levels of the factors chosen. A response surface methodology (RSM), namely a Central Composite Design (CCD), was carried out to analyse the effects. Analysis of variance (ANOVA) was used to choose the significant factors out of the seven factors considered.

### 5.2.1.1 Central Composite Design (CCD)

A Box-Wilson Central Composite Design contains an imbedded factorial or fractional factorial design with center points that is augmented with a group of "star or axial points" that facilitates the estimation of curvature. This allows the user to fit a second order (quadratic) model for the response variable. In statistics, this design is useful in response surface methodology. Additional information is provided in Appendix B.

In general, the CCD design consists of three distinct sets of experimental runs:

1. A factorial or perhaps fractional factorial design involving the factors under study with two levels considered for each factor; i.e. the range of each factor is +/- 1 level.
2. A set of centre points; i.e. experimental runs with the input factors set to the median of the values used in the factorial portion. This point is often replicated in order to improve the precision of the experiment by estimating the pure error.
3. A set of star or axial points ($\alpha$); i.e. experimental runs identical to the centre point runs except for one factor which is assigned values outside the range and both below and above the median of the two factorial levels. All factors

are varied in this way. A CCD design with k factors will always have 2k star points. Usually, in order to maintain rotatability (See Appendix B), the value of $\alpha$ depends on the number of experimental runs in the factorial portion of the central composite design.

The CCD allows approximation of three regression models; linear, interaction (2FI: Two-Factor Interaction) and quadratic terms. The models are defined as follows:

Linear: $$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \varepsilon \tag{5.1}$$

Interaction (2FI): $$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{i<j} \sum_{j=1}^{k} \beta_{ij} x_i x_j + \varepsilon \tag{5.2}$$

Quadratic: $$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{i<j} \sum_{j=1}^{k} \beta_{ij} x_i x_j + \sum_{j=1}^{k} \beta_{jj} x_j^2 + \varepsilon \tag{5.3}$$

where $y$ is the response, $\beta_i$ is the regression coefficient, $\beta_0$ is the overall average, $x_i, x_j$ are variables that represents factor $i$ and $j$ respectively, $k$ is the number of factors, $\beta_{ij}$ represents interaction terms, $\beta_{jj}$ represents pure second order or quadratic effects and $\varepsilon$ is the error estimate.

Table 5.1 shows the DOE factors at the two levels considered. The two levels chosen were based on the nominal values of the commercially available Lumenera camera system used in the lab. The zero level value is the nominal values for each of the camera parameters. Parameters are represented by the letters A, B, C, … , and are therefore referred to as factor A, factor B, factor C, etc.

**Table 5.1: Design space for the study of sum of the square of the error**

| Factors | Level (-) | Level (+) |
|---|---|---|
| A. Effective focal length (mm) | 34.5 | 35.5 |
| B. $U_0$ (pixels) | 686 | 706 |
| C. $V_0$ (pixels) | 510 | 530 |
| D. Rad1, $k_1$ ($mm^{-2}$) | -0.0005 | +0.0005 |
| E. Rad2, $k_2$ ($mm^{-4}$) | $-1 \times 10^{-5}$ | $1 \times 10^{-5}$ |
| F. Tan1, $p_1$ ($mm^{-1}$) | -0.0003 | +0.0003 |
| G. Tan2, $p_2$ ($mm^{-1}$) | -0.0003 | +0.0003 |

### 5.2.1.2 Summary of Experiment Methodology

A series of numerical experiments were conducted using MATLAB. The experimental procedures are outlined below:

*Procedure 1 (using nominal focal length, image center and zero distortion)*

Step 1: Define world coordinates of target points

Step 2: Define rotation matrix

Step 3: Rotate and translate target relative to camera coordinate system

Step 4: Calculate the coordinates of undistorted target points in mm relative to center of image

Step 5: Calculate radial and tangential distortions

Step 6: Calculate coordinates of distorted target points in pixels relative to upper left hand corner of image

*Procedure 2 (using different combinations of levels of focal length, image center and distortion)*

Step 1: Procedure 1 explained above is replicated through several runs for the different levels of factors chosen.

Step 2: The response is calculated by taking the square root of the sum of the squares of the differences between the image coordinates obtained using Procedure1 and Step 1 of Procedure 2.

*Experiments using DESIGN-EXPERT Software:*

- The statistical software Design-Expert 7.1.3 was used to create and analyze the experiment designed to measure the effects of the factors mentioned in Table 5.1.
- Seven factors were each analyzed at 2 levels and one center point using a CCD design. This resulted in 79 runs (half CCD) of the calibration model implemented in MATLAB.
- The 2 levels were chosen based on the nominal values of the Lumenera camera system as provided in the technical data sheet.
- An ANOVA of the significant factors was run to determine the relative importance of each factor to the response.
- No replications were performed since the output results were obtained from a numerical simulation without any sources of error. Since the experimental setup is not

subject to any uncertainties (e.g. measurement error) subsequent replications would yield identical results. In this case the numerical error is assumed to be insignificant.

- The center point in combination with the additional star points facilitates the check for curvature and hence permits the software to fit a second order model.
- Seventy-nine computer simulations were performed in MATLAB and subsequent response values were obtained for each of the 79 treatment combinations.
- The Response Surface Methodology based on Central Composite Design, determined an optimum point on a nonlinear curve between the end point levels.
- Calibration Model Assumptions: Any additional factors other than the seven intrinsic factors under consideration are assumed to be negligible and are not included in the current study; i.e. zero contribution to the response is assumed.

## 5.3 Analysis of Experimental Results and Discussions

The analysis was performed by selecting a half CCD resulting in fewer runs; i.e. 79 runs compared to a full CCD which requires 143 runs. A quadratic model was chosen as the best choice of model that can be fitted from the CCD as compared to the models like linear and 2FI (two factor interaction) by analyzing the fit summary tables (Sequential Model Sum of Squares, Lack of Fit Tests, and Model Summary Statistics) obtained from the Design-Expert software. The experimental results from the Design-Expert software are described in this section.

### 5.3.1 Statistical Test for Significance using ANOVA

Firstly, in order to improve the final (regression) model, model reduction was performed to lessen the number of insignificant model terms (factors) and to determine the best subset of regressor variables (predictor variables or regressors) to include. Backward elimination with $\alpha_s = 0.10$ was used for model reduction (Myers and Montgomery, 2009). The reduced ANOVA (Table 5.2) of the chosen design indicates that the model selected is significant with a model F-Value of 596825.16. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise. In general, if the test statistic F-value >> 1, the corresponding factor effects are considered to be significant. That is, these factors will significantly affect the response variable(s). This significance level approach is done by comparing the calculated F-value to 5% F-table (Percentage Points of the F Distribution) using the degree of freedom (df) of main effects (e.g. A, B, C, ..G) and degree of freedom of error, i.e. residual (refer to Table 5.2). Furthermore, the significance of the effect can also be tested from the computed P-value. If the P-value, "Prob > F" is less than 0.05 (P-value $< \alpha_s = 0.05$), the model effect is said to be highly significant and if "Prob > F" is less than 0.1 (P-value $< \alpha_s = 0.1$), the model effect is said to be significant. Values greater than 0.10 indicate the model terms are not significant. Therefore, several factors including D, AD, AE, BG, CF, $A^2$, $B^2$, $C^2$, $D^2$, $E^2$, $F^2$ and $G^2$ are highly significant model terms, which includes a main effect, four interaction effects, and quadratic effects of all factors considered for the study. Factors like A, B, C, E, F, G

are not significant but must be included in the model to maintain the hierarchy (Myers and Montgomery, 2009).

**Table 5.2: Reduced ANOVA for the calibration model**

**ANOVA for Response Surface Reduced Quadratic Model**

**Analysis of variance table [Partial sum of squares - Type III]**

| Source | Sum of Squares | df | Mean Square | F Value | p-value Prob > F | |
|---|---|---|---|---|---|---|
| Model | 70.63 | 18 | 3.92 | 5.968E+005 | < 0.0001 | significant |
| *A-Focal Length* | *1.317E-005* | *1* | *1.317E-005* | *2.00* | *0.1622* | |
| *B-U0* | *0.000* | *1* | *0.000* | *0.000* | *1.0000* | |
| *C-V0* | *0.000* | *1* | *0.000* | *0.000* | *1.0000* | |
| *D-Rad1* | *9.442E-005* | *1* | *9.442E-005* | *14.36* | *0.0004* | |
| *E-Rad2* | *8.406E-006* | *1* | *8.406E-006* | *1.28* | *0.2627* | |
| *F-Tan1* | *0.000* | *1* | *0.000* | *0.000* | *1.0000* | |
| *G-Tan2* | *0.000* | *1* | *0.000* | *0.000* | *1.0000* | |
| *AD* | *0.025* | *1* | *0.025* | *3803.01* | *< 0.0001* | |
| *AE* | *2.676E-004* | *1* | *2.676E-004* | *40.70* | *< 0.0001* | |
| *BG* | *0.091* | *1* | *0.091* | *13814.28* | *< 0.0001* | |
| *CF* | *0.051* | *1* | *0.051* | *7738.59* | *< 0.0001* | |
| $A^2$ | *8.44* | *1* | *8.44* | *1.284E+006* | *< 0.0001* | |
| $B^2$ | *21.92* | *1* | *21.92* | *3.334E+006* | *< 0.0001* | |
| $C^2$ | *21.92* | *1* | *21.92* | *3.334E+006* | *< 0.0001* | |
| $D^2$ | *0.059* | *1* | *0.059* | *9043.69* | *< 0.0001* | |
| $E^2$ | *0.63* | *1* | *0.63* | *96354.33* | *< 0.0001* | |
| $F^2$ | *0.028* | *1* | *0.028* | *4212.21* | *< 0.0001* | |
| $G^2$ | *1.594E-003* | *1* | *1.594E-003* | *242.48* | *< 0.0001* | |
| Residual | 3.879E-004 | 59 | 6.574E-006 | | | |
| Cor Total | 70.63 | 77 | | | | |

The interactions between the various factors that were found to be significant are:

1. "AD" which represents the interaction between focal length (factor "A") and radial distortion coefficient 1 (factor "D").
2. "AE" which represents the interaction between focal length and radial distortion coefficient 2 (factor "E").
3. "BG" which represents the interaction between X image center, $U_0$ (factor "B"), and tangential distortion coefficient 2 (factor "G").
4. "CF" which represents the interaction between Y image center, $V_0$ (factor "C") and tangential distortion coefficient 1 (factor "F").
5. Finally, the quadratic effects of all seven intrinsic parameters were found to be significant as demonstrated by the seven squared terms.

Table 5.3 shows the summary statistics of the model undertaken.

**Table 5.3: Model summary statistics**

| | | | |
|---|---|---|---|
| Std. Dev. | 2.564E-003 | R-Squared | 1.0000 |
| Mean | 3.50 | Adj R-Squared | 1.0000 |
| C.V. % | 0.073 | Pred R-Squared | 1.0000 |
| PRESS | 6.831E-004 | Adeq Precision | 2938.622 |

"Std. Dev." is the square root of the error mean square (Montgomery, 2008). "C.V." is the coefficient of variation that measures the unexplained or residual variability in the data as a percentage of the mean of the response variable. "PRESS" stands for "prediction error sum of squares," and it is a measure of how well the model for the experiment is likely to predict the response in a new experiment. Small values of PRESS

are desirable. The quantity "R-squared" is interpreted as the proportion of the variablity in the data explained by the ANOVA model. The "Pred R-Squared" of 1.0000 is in reasonable agreement with the "Adj R-Squared" of 1.0000. "Adeq Precision" measures the signal to noise ratio. Large values of this quantity are desirable, and values that exceed four usually indicate that the model will give reasonable performance in prediction. In this case the ratio of 2938.622 indicates an adequate signal, and, therefore, this model can be used to navigate the design space.

In designed experiments, "Model adequacy checking" is an important part of the data analysis procedure. It is also useful in building regression models. The model diagnostic plots (e.g. residual plots) obtained from the designed experiment are always examined for a regression model (Montgomery, 2008). In general, when building a regression model from DOE, it is always necessary to (1) examine the fitted model to ensure that it provides an adequate approximation to the true system and (2) verify that none of the least squares regression assumptions (ANOVA) are violated. Specifically, these assumptions are that the observations are adequately described by the model,

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \tag{5.4}$$

and that the errors (or residuals) represented by,

$$e_{ij} = y_{ij} - \hat{y}_{ij} \tag{5.5}$$

are normally and independently distributed with mean zero and constant variance $\sigma^2$, where $y_{ij}$ is the $ij$th observation, $\mu_i$ is the mean of the $i$th factor level or treatment, $\tau_i$ is

the $i$th treatment effect, $\varepsilon_{ij}$ is a random error component that incorporates all other sources of variability in the experiment, and $\hat{y}_{ij}$ is an estimate of the corresponding observation $y_{ij}$. Violations of the basic assumptions and many types of model inadequacies can be easily investigated by the examination of residuals, especially by the graphical analysis of the residuals. If the model is adequate, the residuals should contain no obvious patterns; i.e. it should be structureless. See Appendix C for more information regarding how the model adequacy checking was done through the graphical analysis of residuals including the model diagnostic plots.

### 5.3.2 Regression Analysis

Regression analysis in designed experiments is a general approach used to fit an empirical model (also named as prediction models); an interpolation equation developed for the response variable in the experiment. Regression models are used to fit more precise models rather than giving a rough approximation of the model. Equations (5.6) and (5.7) are the final model equations, i.e. fitted regression models (a quadratic model fit) obtained in terms of coded and actual factors, respectively. These models are second order response surface models with seven independent or regressor variables.

**Final Equation in Terms of Coded Factors:**

$$
\begin{aligned}
\text{Sqrt(Sum Square of error)} = {} & +0.86 \\
& +4.360\text{E-}004 * A \\
& +5.319\text{E-}007 * B \\
& -5.721\text{E-}007 * C \\
& -1.167\text{E-}003 * D \\
& -3.483\text{E-}004 * E \\
& +1.954\text{E-}007 * F \\
& -9.162\text{E-}007 * G \\
& -0.020 * A * D \\
& -2.045\text{E-}003 * A * E \\
& -0.038 * B * G \\
& -0.028 * C * F \\
& +0.75 * A^2 \\
& +1.20 * B^2 \\
& +1.20 * C^2 \\
& +0.063 * D^2 \\
& -0.20 * E^2 \\
& -0.043 * F^2 \\
& +0.010 * G^2
\end{aligned}
\tag{5.6}
$$

**Final Equation in Terms of Actual Factors:**

Sqrt(Sum Square of error) =

+12719.99179

-208.71537 * Focal Length

-16.72030 * U0

-12.49218 * V0

+2764.80612 * Rad1

+14278.67108 * Rad2

+4887.11606 * Tan1

+8739.60120 * Tan2

-79.06117 * Focal Length * Rad1

-408.95717 * Focal Length * Rad2

-12.55690 * U0 * Tan2

-9.39830 * V0 * Tan1

+2.98166 * Focal Length$^2$

+0.012012 * U0$^2$

+0.012012 * V0$^2$

+2.50224E+005 * Rad1$^2$

-2.04188E+009 * Rad2$^2$

-4.74360E+005 * Tan1$^2$

+1.13812E+005 * Tan2$^2$

(5.7)

## 5.3.3 Interpretation of the Results

The model graphs representing the significant interactions between the various factors are shown below in the form of curvatures.

*Interaction between Focal Length and Radial Distortion Coefficient 1:*

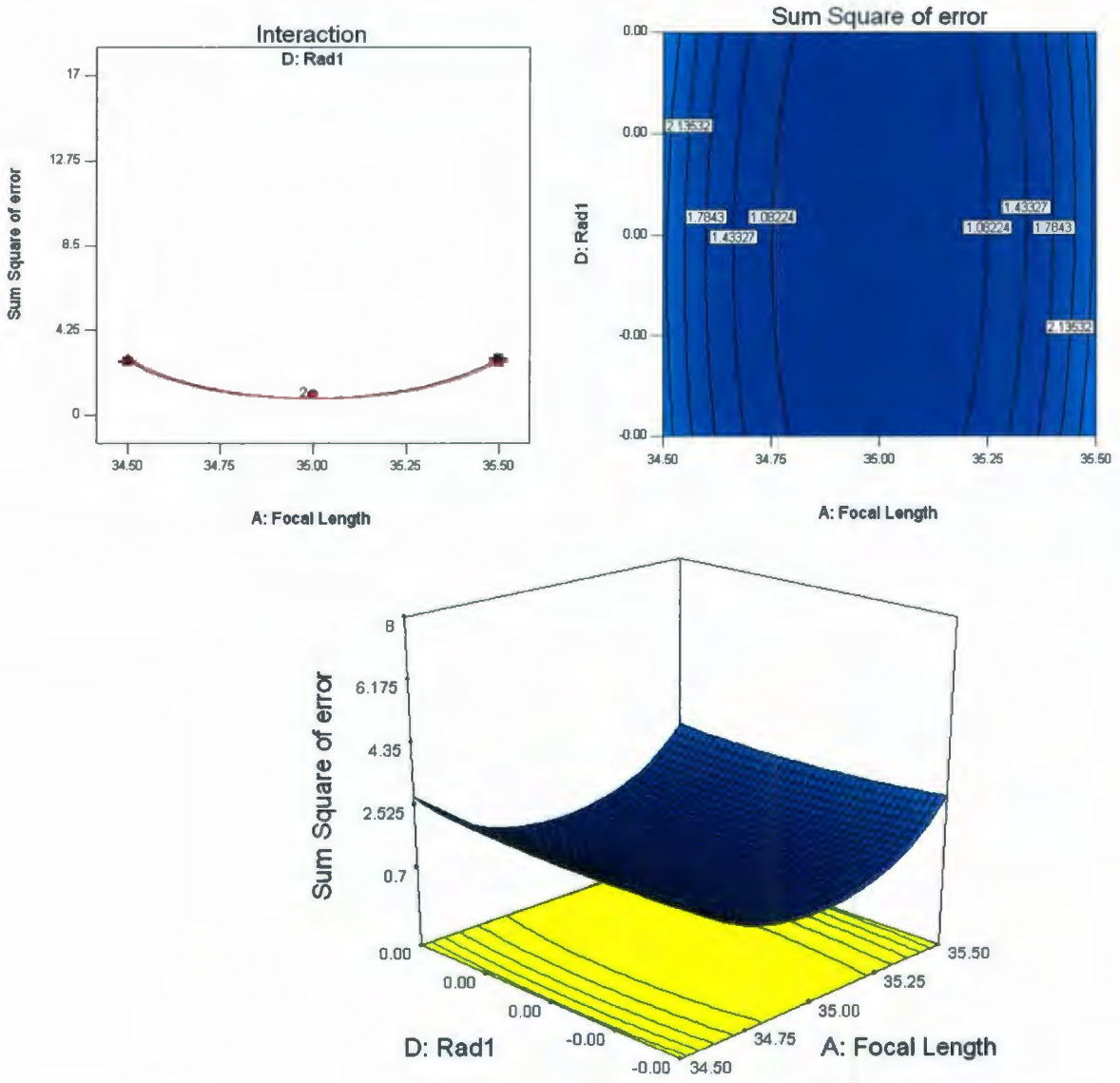



Figure 5.1: Interaction between factors "A" and "D"

The interaction between factors "A" and "D" is shown in the above plots (Figure 5.1). As expected, the sum of the square of the error is minimum when the focal length (factor

"A") and radial distortion coefficient 1 (factor "D") are at their nominal value; however, it can be seen that for a range of values for "D" the error is almost constant.

*Interaction between Focal length and Radial Distortion Coefficient 2:*

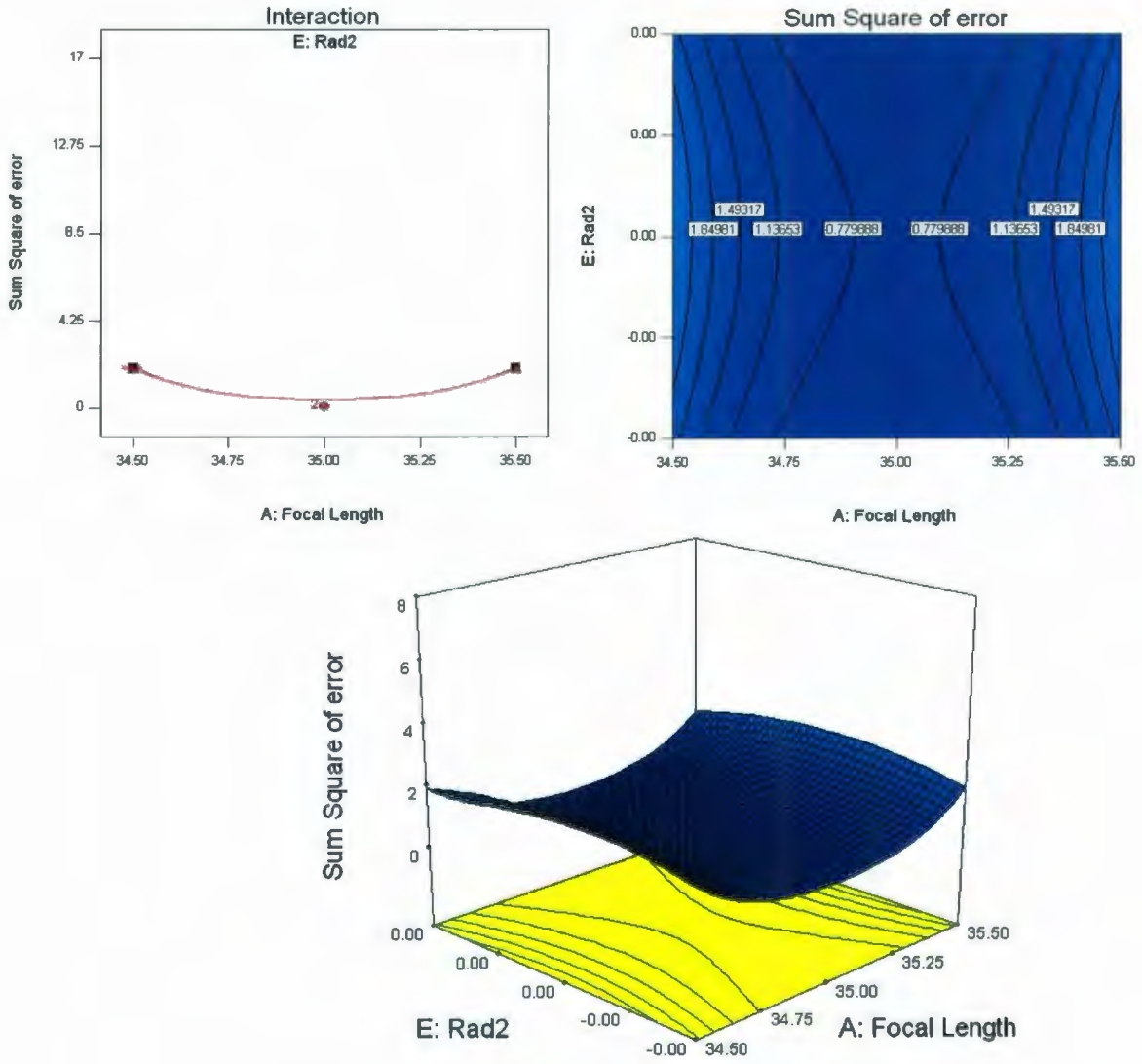
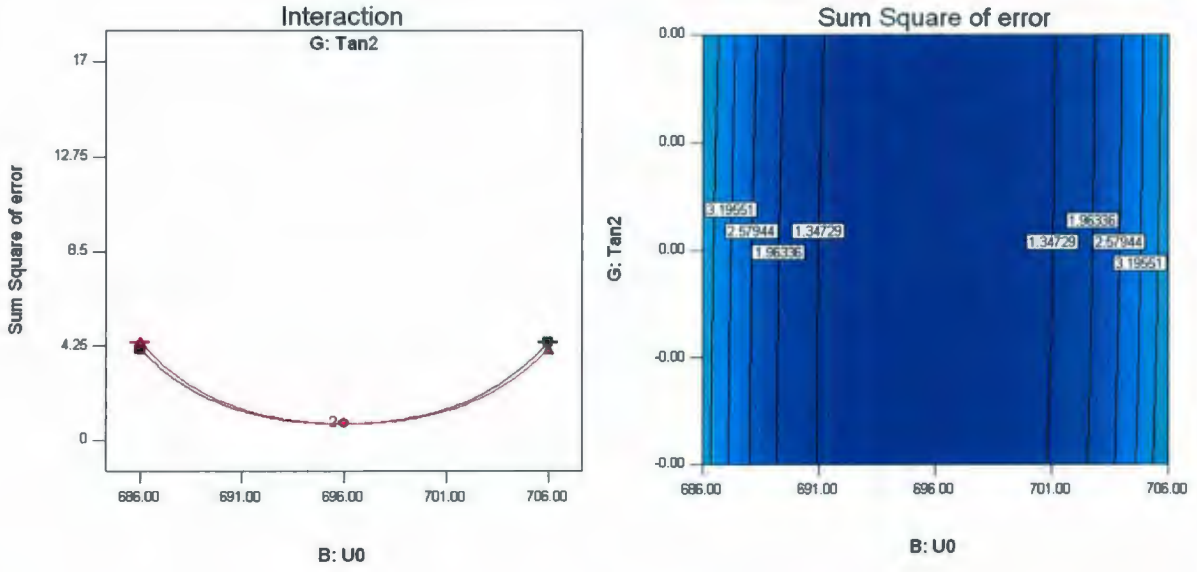



Figure 5.2: Interaction between factors "A" and "E"

The above plots, Figure 5.2, represent the interaction between factors "A" and "E". It can be seen that it is difficult to identify a minimum value for the error which could potentially make it difficult to find the actual camera parameters resulting in an inaccurate camera calibration.

*Interaction between Image Centre, $U_0$, and Tangential Distortion Coefficient 2:*



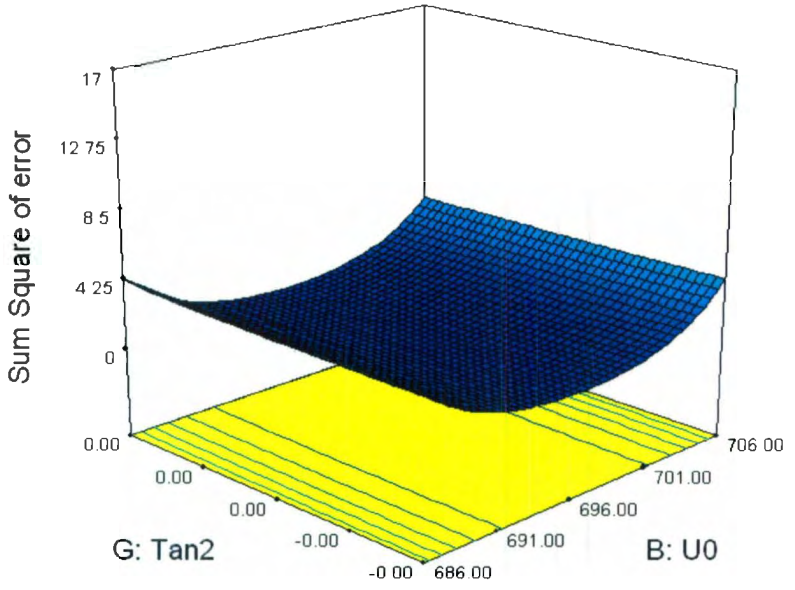



Figure 5.3: Interaction between factors "B" and "G"

The interaction between factors "B" and "G" can be seen from the above plots (Figure 5.3). The sum of the square of the error is minimum when the X image center, $U_0$ (factor "B") and tangential distortion coefficient 2 (factor "G") are at their nominal value, but it can be seen that for a range of values for "E" the error is almost constant.

*Interaction between Image Centre, $V_0$, and Tangential Distortion Coefficient 1:*

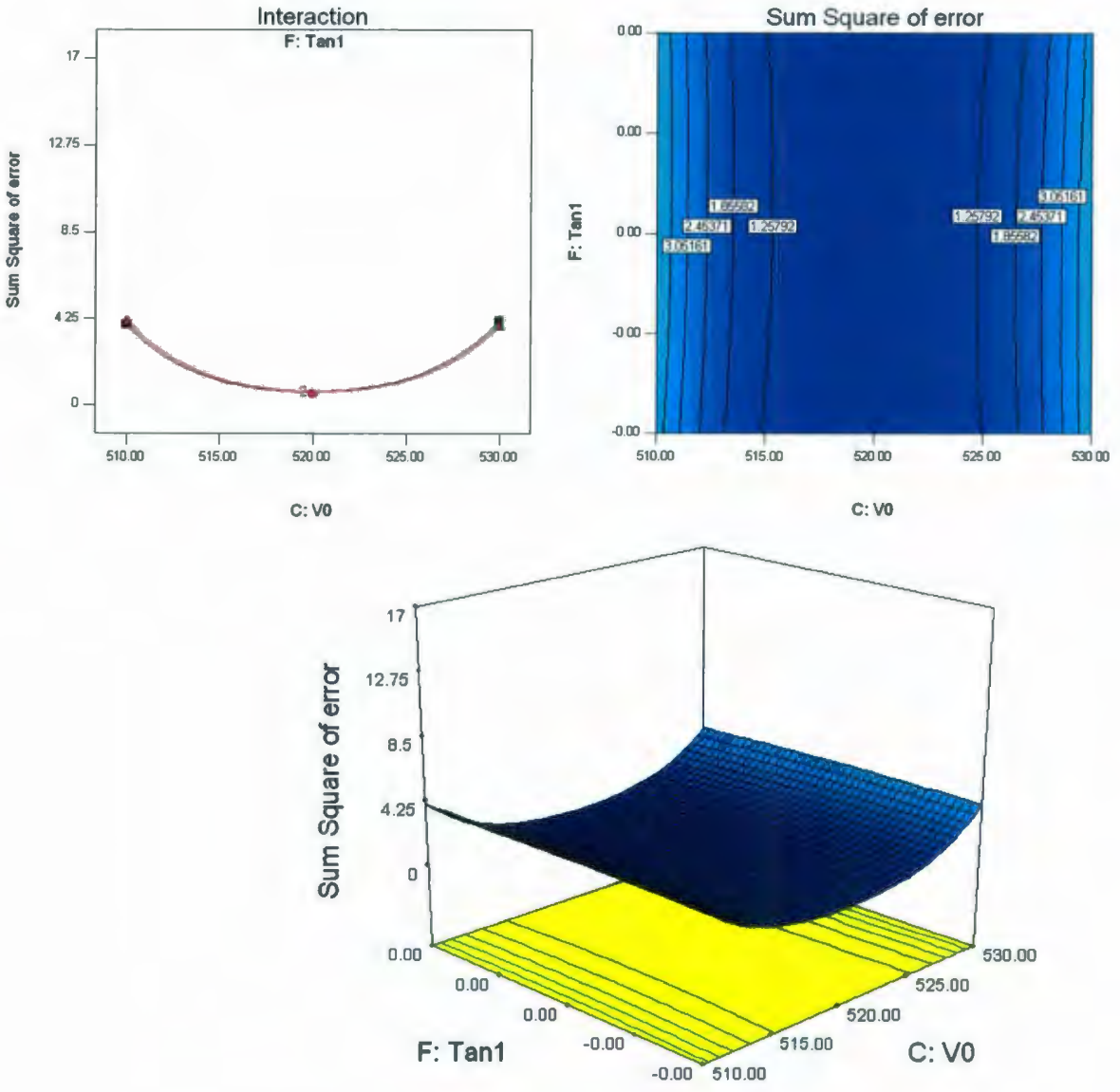



Figure 5.4: Interaction between factors "C" and "F"

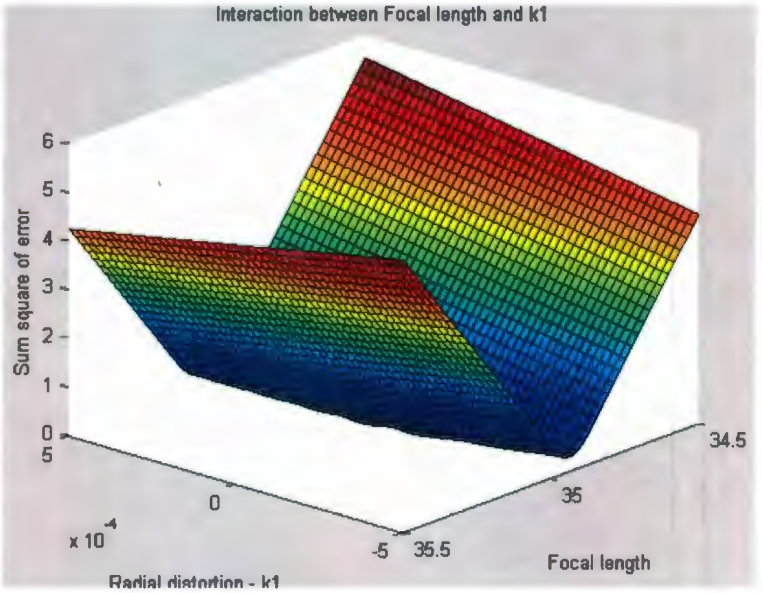
The above plots, Figure 5.4, represent the interaction between factors "C" and "F". The sum of the square of the error is minimum when the Y image center, $V_0$, (factor "C") and

the tangential distortion coefficient 1 (factor "F") are at their nominal value, but it can be seen that for a range of values for "E" the error is almost constant.

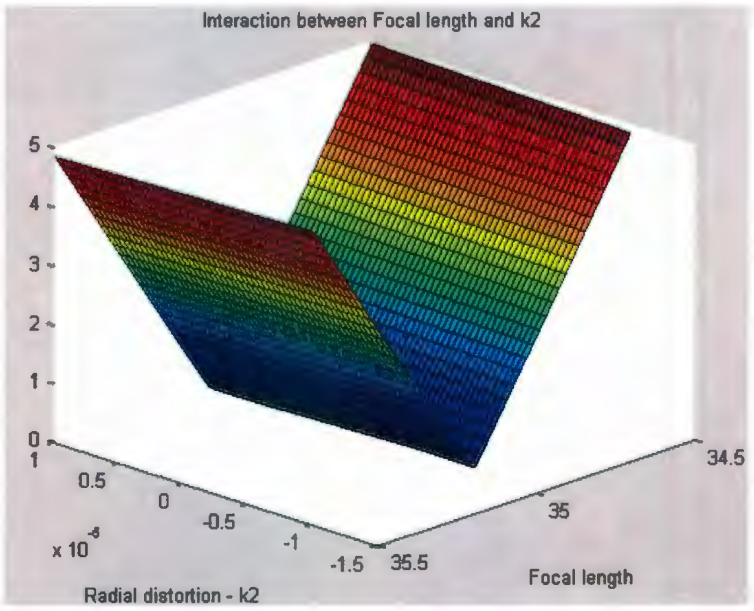
The results of the design of experiments study indicate that due to the interactions between parameters the calibration algorithm will not necessarily converge to the minimum value when the data is subject to measurement error and the error cannot necessarily be attributed to a single camera parameter. The above interaction plots demonstrate that the error is potentially due to the combined effects of multiple parameters that cannot be distinguished separately. This makes it difficult to get accurate calibration results in practice, although the techniques and algorithms are theoretically sound.

## 5.4 Validation of the Model

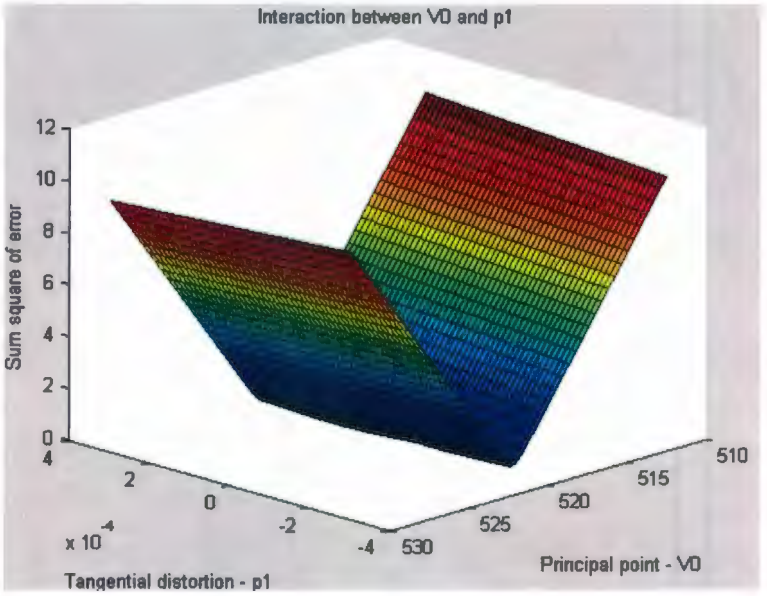
Validation of the model is evaluated by comparing the 3-D plots obtained using Design-Expert with a more detailed simulation implemented in MATLAB. For this purpose, 50 points were considered between the high and low level values of factors that were found to have strong interactions. The error for each combination of these 50 points was plotted. The interaction between "AD", "AE", "BG" and "CF" were obtained and are shown below (Figure 5.5):


Interaction between Focal length and k1
Sum square of error
6
5
4
3
2
1
0
5
0
-5
x 10^-4
Radial distortion - k1
35.5
35
34.5
Focal length

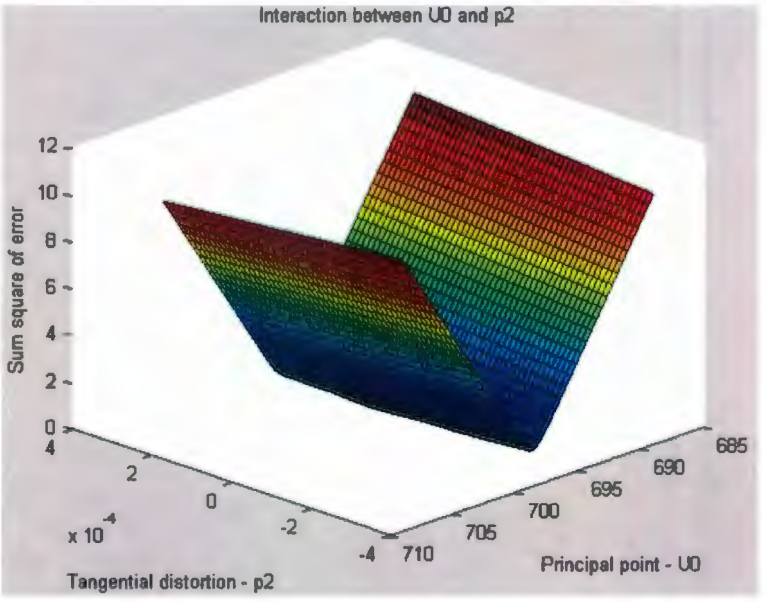
(a)


Interaction between Focal length and k2
5
4
3
2
1
0
1
0.5
0
-0.5
-1
-1.5
x 10^-5
Radial distortion - k2
35.5
35
34.5
Focal length

(b)




(c)




(d)

Figure 5.5: Plots of interaction between factors generated in MATLAB (a) Interaction between "A" and "D", (b) Interaction between "A" and "E", (c) Interaction between "B" and "G", (d) Interaction between "C" and "F"

The plots shown above (Figure 5.5) match the 3-D surface plot obtained from the Design-Expert software. The error is minimum (almost zero) when the factors are at their nominal values; i.e. at their center values, for Figure 5.5 (a), (c) and (d). When the focal length, X image centre, and Y image center are in the vicinity of their nominal values, there is little variation in the error for a relatively wide range of values of radial distortion coefficient 1, tangential distortion coefficient 2, and tangential distortion coefficient 1, respectively. For instance, small variations in focal length and radial distortion coefficient 1 result in no appreciable change in response (i.e. error) which is a prerequisite for accurate camera calibration. This makes experimental camera calibration very difficult to perform in practice. The same conclusion can be drawn from the relationship between image centre and tangential distortion.

In the case of the interaction graph between focal length and radial distortion coefficient 2, the result suggests that it is difficult to estimate the minimum value for error at the nominal values of these factors. Hence the radial distortion coefficient 2 is effectively negligible and can be removed from the distortion model of the camera; however, the ANOVA analysis in the DOE shows it as a significant interaction factor; a significant finding that goes against much of the published camera calibration literature.

To conclude, in vision based applications where high accuracy is a requirement, the various interactions between the intrinsic parameters play a critical role that can significantly affect the results of the camera calibration procedure. One of the most significant findings from this study is that certain camera parameters can effectively be removed from the calibration model with no discernible loss in accuracy.

# Chapter 6

# A New Approach to Camera Calibration

As discovered from the previous studies, it can be said that for 3-D object analysis, *accuracy* is an important aspect that needs to be taken into consideration when performing geometric camera calibration. Inaccurate calibration results will adversely affect the estimation of the position and orientation of 3-D objects as well as 2-D image analysis in machine vision applications. Furthermore, the DOE study concluded that in vision based applications where high accuracy is a requirement, the various interactions between the intrinsic camera parameters play a crucial role that can significantly affect the results of the camera calibration procedure. It was found that certain intrinsic parameters like focal length and image center are strongly coupled with radial and tangential distortions, respectively; i.e. highly significant interaction effects were

revealed. It was noticed from the interaction plots that for a range of values of these parameters there was no appreciable change in the error response; i.e. error was almost constant which renders it difficult to distinguish the error causing treatments or factor effects (i.e. parameters) seperately. This would make experimental camera calibration very challenging to perform in practice. Based on these DOE results, the final focus of this thesis is to propose a new approach towards geometric camera calibration in which image center and focal length are calculated independently of the lens distortion. The idea behind this study is to demarcate certain intrinsic parameters like image center and focal length and compute them separately and precisely without taking into consideration lens distortion. A practical experiment was conducted for this study. The experimental results show that this new technique for finding the image center and focal length of a camera has the potential to improve the calibration accuracy. An elaborated description of the proposed technique is presented in the following sections.

## 6.1 Overview of the Experimental Setup

The experimental setup includes a calibration test rig (consisting of a precision X-Y table, solid-state camera, LED lighting, stepper motor driven ball screws for controlling the table motion) and a microcontroller interface to a PC. The structure of the test rig used for calibration purposes is shown in Figure 6.1. The table is made up of aluminum extrusions. It incorporates a camera (the camera system used here is a Lumenera Lm135 1.4 megapixel industrial USB 2.0 camera) placed at the top of the test rig looking

downwards at the X-Y table. The X-Y table is driven by two stepper motors with a stepping resolution of 0.0005 inches or 0.0127mm. The movement of the table was aligned to ensure perpendicularity between the two axes. The maximum positional error of the table was experimentally determined to be 25 microns.

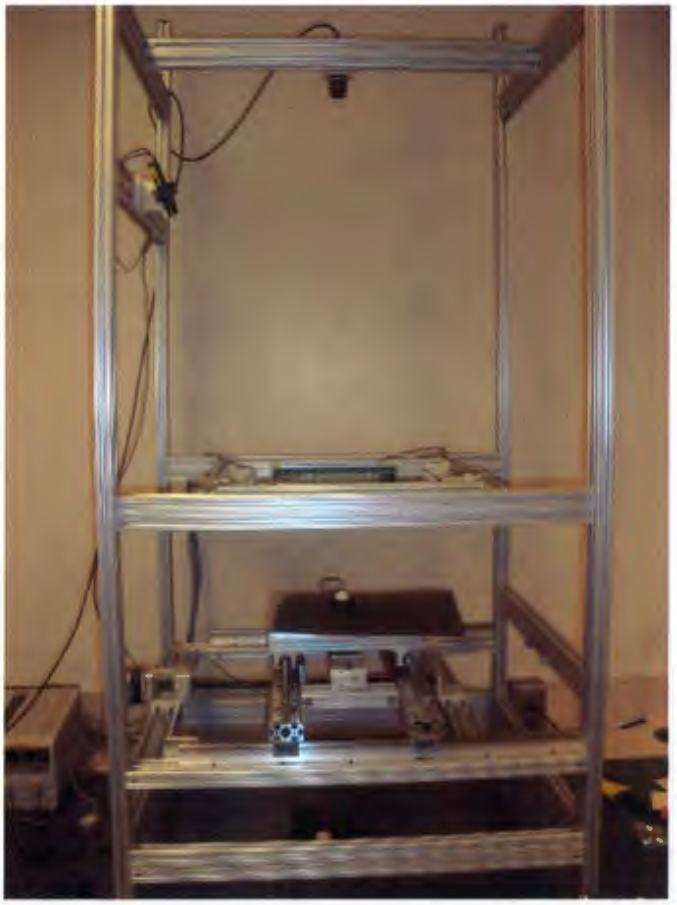


Figure 6.1: Test rig for calibration

The PIC based microcontroller interface used in this study serves as an interface circuit between the test rig and the PC. Figure 6.2 represents the microcontroller interfacing circuit.

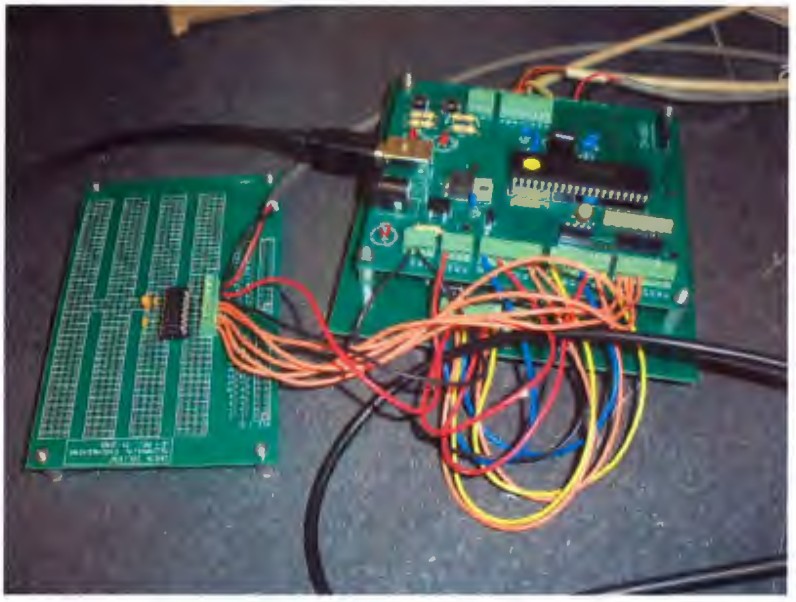


Figure 6.2: PIC based microcontroller interface

The specific role of the microcontroller interface is to precisely control the movement of the stepper motor driven X-Y table on the test rig. The two motors rotate either clockwise or counter-clockwise in accordance with output signals from the microcontroller interface circuit. Figure 6.3 shows the connection between the microcontroller interface and one stepper motor including a power supply.

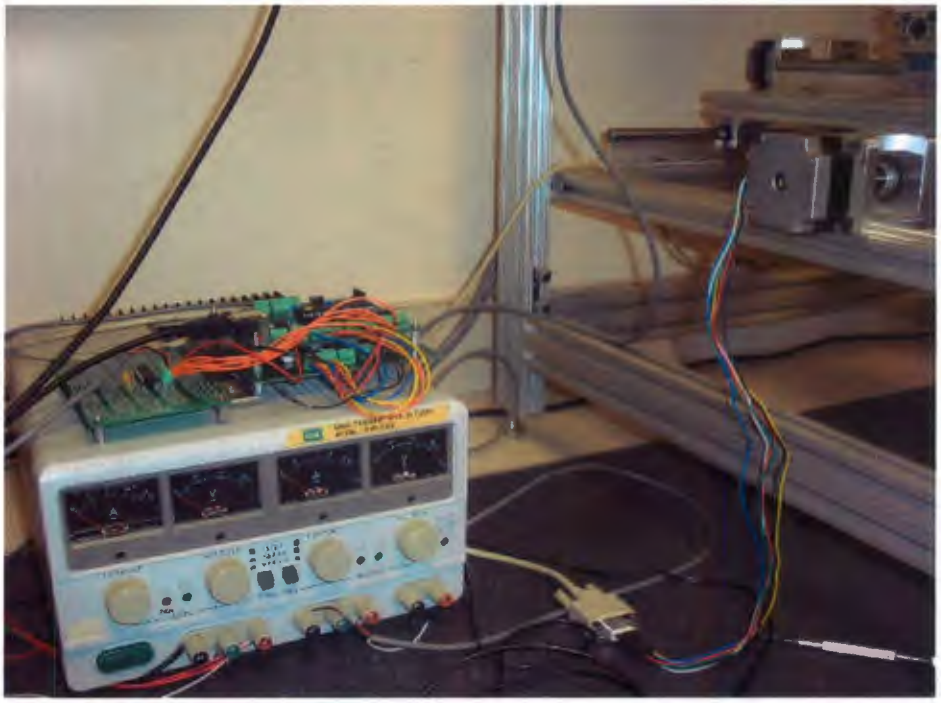
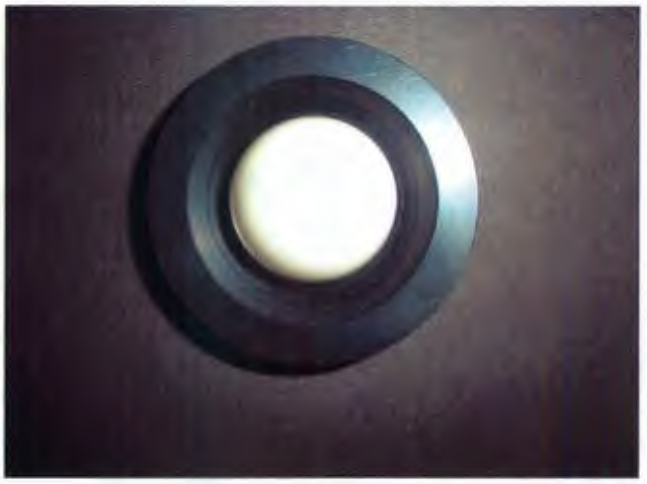


Figure 6.3: PIC based microcontroller interface with a stepper motor

For the proposed technique, a computer application including a Graphical User Interface (GUI) was developed in MATLAB. This application also serves as an interface between the PC and the camera. The microcontroller interface moves the table by means of stepper motor driven ball screws and acquires the image of the calibration target at predetermined grid positions with the help of the image acquisition software developed for the camera. Images are stored in the folder specified by the user for further processing.

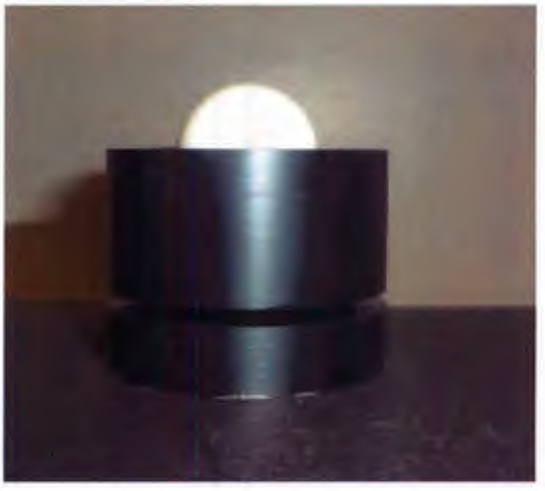
A precisely ground (3-D) spherical target with a sphericity of 2.5microns , placed at two different height levels was used to obtain the control points for the calibration procedure. Figure 6.4 shows the 3-D spherical target used for this work. The advantages of using a 3-D spherical target are: 1) Being isotropic in nature, it will simplify the

process of calibration greatly; 2) More accurate and efficient for high accuracy applications than 2-D planar target; 3) Suitable geometric and optical/radiometric properties; 4) More target and image point locations can be generated within a small area.



(a)

(b)



(c)

Figure 6.4: Precisely ground (3-D) spherical target
(a) Top view of the sphere kept on a cylindrical base (b) Side view of the sphere kept on a cylindrical base (c) Side view of the sphere kept on a cylindrical base of longer height

Four arrays of white LEDs (Light emitting diodes) placed at the four sides of the test rig are used to provide appropriate lighting. Each array consists of 70 discrete LEDs in 5 rows of 14. A power supply is used to supply the required voltage to the LEDs Figure 6.5.

(a) (b)

Figure 6.5: LED Lighting
(a) An array of white LEDs (b) Four arrays of white LEDs placed above the target

Further scripts developed in MATLAB were used to implement the proposed calibration technique. These scripts include image processing algorithms to extract the image points with subpixel accuracy, numerical methods like Newton's method for solving nonlinear system of equations, nonlinear least square analysis to fit a nonlinear model to a set of experimental data points, etc. These techniques will be explained in detail in the next section (Section 6.2).

## 6.2 Description of the Proposed Technique

### 6.2.1 Methodology Overview

As mentioned earlier, the approach undertaken for this study was to compute the image center and focal length of the camera independent of lens distortion. Image center, also called the principal point, is the point where the optical axis of the lens pierces the image plane. Focal length is the perpendicular distance between the image plane and the projective center of the lens. The goal of the proposed approach was to find these parameters separately and accurately. The proposed technique employed a precision ceramic sphere (3-D object) as a calibration object. For this, an experimental set up represented by the diagram shown in Figure 6.6 was used to determine the two intrinsic parameters. The diagram represents only a part of the entire experimental set-up neglecting the microcontroller interface circuit and all other related components such as the test rig frame, stepper motors, LEDs, power supply, etc.

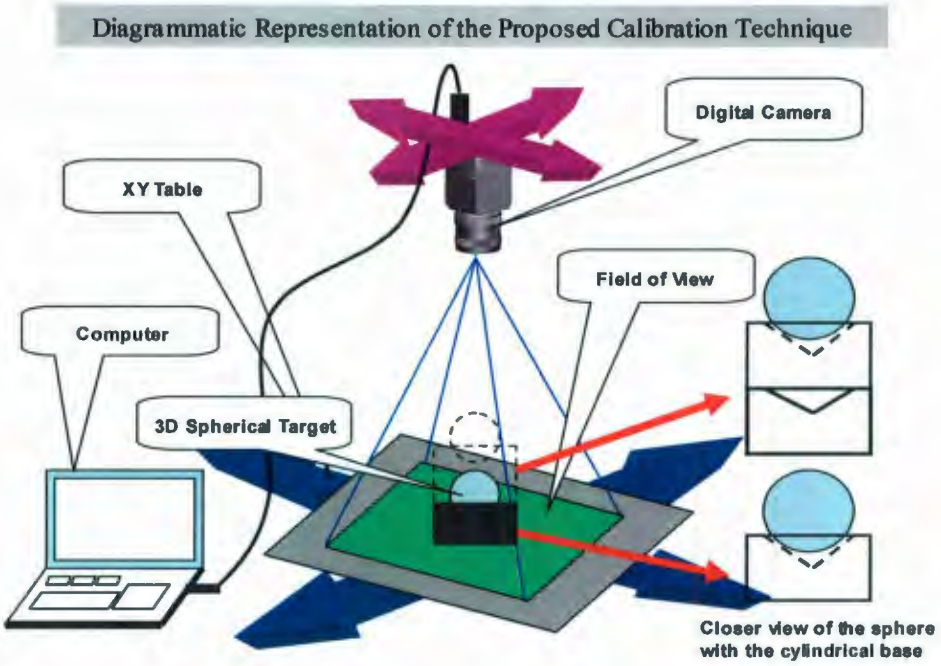



Figure 6.6: Diagrammatic representation of the proposed calibration technique

As shown in the figure, the camera, which is fixed to the top of the calibration test frame, looks down at the spherical target which is mounted on the X-Y table. The sphere is placed in a conically shaped cavity situated at the center of the cylindrical base (shown as black in the diagram). The sphere can be positioned at a second known height relative to the table by introducing a precisely machined spacer block as shown in Figure 6.6 (also refer to Figure 6.4(c)). The cylindrical base and spacer block ensure that the line between the centers of the sphere in the two positions is perpendicular to the table.

A sphere is a non-degenerate quadric surface whose projection (i.e. the occluding contour) is given by the projection of its contour generator which is the conic formed by the tangent rays traveling from the camera center (Wijewickrema et al., 2006). The

contour generator of a sphere is essentially a circle, as any plane through the sphere would result in a circle. In general, the occluding contour of a sphere projects to a conic, which is a circle or an ellipse in the image. There are several reports in the literature which deal with camera calibration using spheres as calibration targets (e.g. Shivaram and Seetharaman, 1998; Agrawal and Davis, 2003; Zhang et al., 2005, 2007; Ying and Zha, 2006). In this study, the occluding contour of the sphere is interpreted as a circle in the image. The image of the sphere is acquired for each position in the table and various image processing algorithms are employed to obtain the 2-D measurement point; i.e. the center of the observed circle. These points are later used to solve for the image center and focal length as described in detail in Section 6.2.4.

### 6.2.2 Processing of the Images

A typical image of the calibration sphere as captured by the camera is shown in Figure 6.7. The four arrays of white LEDs positioned around the table ensure uniform lighting over the field of view of the camera. This renders the object clearer and improves the contrast of the sphere relative to the dark background. The image is processed using a series of image processing algorithms in order to determine the 2-D coordinates of the center of the observed sphere. To estimate the center, it is necessary to first extract the contour of the sphere.

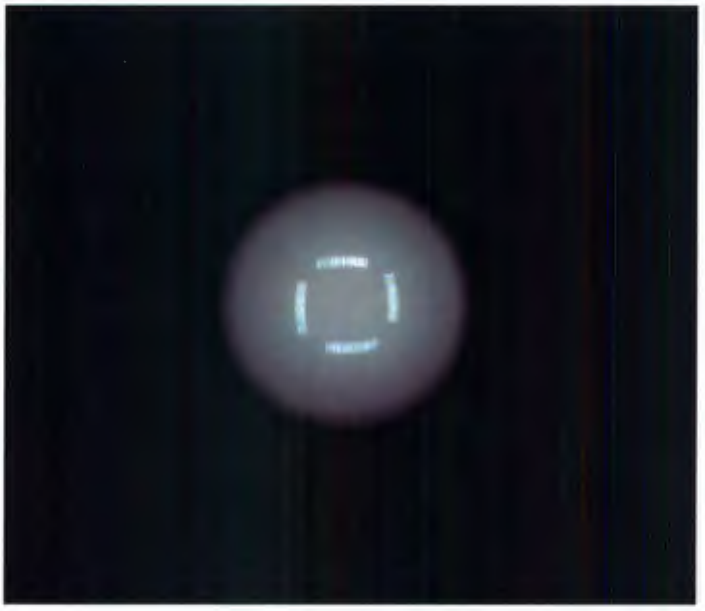


Figure 6.7: A typical image of the calibration sphere

The various steps involved in extracting the contour and finding the center of the image with subpixel accuracy can be described as follows:

1) The acquired RGB image is first converted to a gray scale (or monochrome) image (Figure 6.8). A grayscale image is an image whose pixel values specify intensity values that are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest intensity. Pixel values in an 8 bit grayscale image define gray levels ranging from 0 to 255 (256 gray levels).

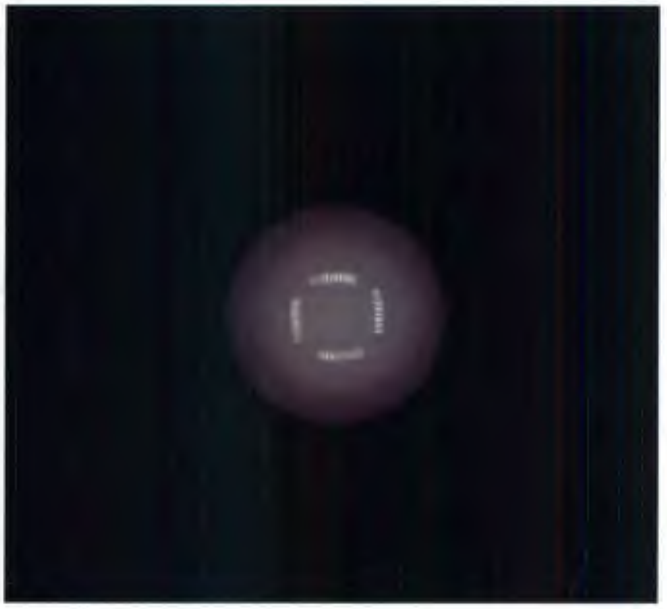


Figure 6.8: A gray scale image of the calibration sphere

2) Thresholding and converting to a binary image: Thresholding is a process of image segmentation. Thresholding creates a binary image from a grayscale image by converting all pixels below some particular threshold to zero and all pixels above that threshold to one. The automatic thresholding algorithm (Otsu's method) was used to find this optimum threshold and thereby convert the grayscale image to a binary image. A binary image is required to apply the contour tracking algorithm (Step 3). Otsu's algorithm assumes that the image to be thresholded contains two groups of pixels; i.e. foreground and background. The optimum threshold separating these two groups is chosen such that the sum of the weighted within-group variances is minimum. This can be expressed mathematically as (Morse, 1998-2000 [Online]):

$$\sigma^2_{Within}(T) = n_B(T)\sigma^2_B(T) + n_F(T)\sigma^2_F(T) \tag{6.1}$$

where $\sigma^2_{Within}(T)$ = Within-group variances at threshold 'T'

$n_B(T)$ = Total number of pixels in the group 'Background' (below threshold)

$n_F(T)$ = Total number of pixels in the group 'Foreground' (above threshold)

$\sigma^2_B(T)$ = The variance of the pixels in the group 'Background'

$\sigma^2_F(T)$ = The variance of the pixels in the group 'Foreground'

Obtaining the minimum within-group variance is equivalent to selecting a threshold that maximizes the squared difference between the group means which is referred to as the between-group variance. This can be expressed as:

$$\sigma^2_{Between}(T) = n_B(T)n_F(T)[\mu_B(T) - \mu_F(T)]^2 \tag{6.2}$$

where $\sigma^2_{Between}(T)$ = Between-group variances at threshold 'T'

$\mu_B(T)$ = The mean of the pixels in the group 'Background'

$\mu_F(T)$ = The mean of the pixels in the group 'Foreground'

The implementation of Otsu's Method used here is based on the concept of within-group variance works as follows:

- Compute the histogram of the gray scale image to identify the number of pixels at each particular gray level and classify them into two groups.
- Find the variance of the pixels of each group.
- Multiply the variance with the corresponding number of pixels in each group.
- Sum the products to obtain the within-group variances.
- The optimum threshold is the one that minimizes the within-group variances.

3) Contour tracking / Border following: Once the binary image is obtained, contour tracking algorithm is applied to identify the pixels that fall on the boundary of the object (sphere). Boundary pixels are defined as those pixels that have a neighbor that belongs to the background region. There are two standard code definitions used to represent boundaries: code definitions based on 4-connectivity (crack code) and code definitions based on 8-connectivity (chain code) (Krouglicof, [Online]). In this case, crack code is used to represent boundaries. The contour tracking algorithm for generating crack code is as follows:

- Identify a pixel 'P' that belongs to the class 'Objects' (i.e. P=1) and a neighboring pixel (4 neighbor connectivity) 'Q' that belongs to the class 'Background' (i.e. Q=0)
- Depending on the relative position of 'Q' relative to 'P', identify pixels 'U' and 'V' as follows:

| CODE 0 | | CODE 1 | | CODE 2 | | CODE 3 | |
|---|---|---|---|---|---|---|---|
| V | Q | Q | P | P | U | U | V |
| U | P | V | U | Q | V | P | Q |

- Pixels 'U' and 'V' are used to determine the next move; i.e. the next element of crack code is summarized in the following truth table:

| U | V | P' | Q' | TURN | CODE |
|---|---|---|---|---|---|
| X | 1 | V | Q | RIGHT | CODE-1 |
| 1 | 0 | U | V | NONE | CODE |
| 0 | 0 | P | U | LEFT | CODE+1 |

- The contour tracking algorithm terminates when the point 'P' returns to the starting point.

4) Object Recognition using Blob Analysis: Once the image has been segmented into regions or classes representing the objects in the image, the next step is to generate a high level description of the various objects. In the present case, there is only one object of interest within the image. Object recognition is often done with the help of a set of form parameters that describe the object. Ideally these parameters are invariant with respect to position, orientation and scale; e.g. number of holes in the object, compactness or

complexity (perimeter$^2$/area), and moment invariants. The generalized moment equation for a digital image of size $n$ by $m$ pixels can be expressed as:

$$M_{ij} = \sum_{x=1}^{n} \sum_{y=1}^{m} x^i y^j f(x,y) \tag{6.3}$$

where $M_{ij}$ represents the $(i,j)^{th}$ moment. In the case of a binary image, the function $f(x,y)$ takes a value of '1' for pixels belonging to the class 'object' and '0' for the class 'background'. Note that the generalized moments can be computed directly from the crack code during contour tracking.

The first three moments; i.e. $M_{00}$, $M_{01}$ and $M_{10}$, are particularly useful for determining the center of mass or centroid of the object of interest (i.e. image of the sphere). The coordinates of the centroid can be estimated from the moments as follows:

$$(\bar{X},\bar{Y}) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}}\right) \tag{6.4}$$

where $M_{00}$ = The total number of pixels in the object; i.e. the "Area" of the object

$M_{10}$ = The sum of the x-coordinate values for each pixel in the object

$M_{01}$ = The sum of the y-coordinate values for each pixel in the object

The radius, $R$, of the image of the sphere can be determined from the area, $M_{00}$, using the following expression:

$$R = \sqrt{M_{00}/\Pi} \tag{6.5}$$

The centroid and radius obtained above provide approximate values for the extracted image of the sphere; i.e. without subpixel accuracy. The next step is to identify the exact values for these parameters in order to locate the exact measurement point in the image.

5) Circles Fitting: Fit a circle to the image of the sphere using the estimated centroid and radius. Fit two additional circles in the vicinity of this fitted circle using a radius above (radius_maximum) and below (radius_minimum) the estimated radius. Figure 6.9 shows the resulting image.

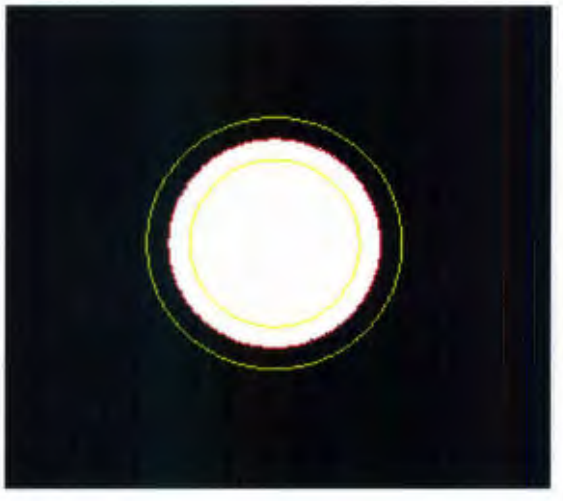


Figure 6.9: Three circles fitted on the binary image of the calibration sphere

6) Edge Detection: The next step in the processing of the image is to apply an edge detection algorithm to identify the points in the digital image at which the image intensity

changes sharply; i.e. a jump in intensity from one pixel to the next. Edges characterize boundaries and are areas with strong intensity contrast (Green, 2002). In many applications, detecting the edges in an image significantly reduces the amount of data, filters out useless information, and preserves the important structural properties in an image. There are many ways to perform edge detection. Here, a gradient method of edge detection; namely Sobel method, was used. It operates by applying a 2-D spatial gradient operator on an input grayscale image. It can also provide directional information; i.e. the direction of the maximum intensity gradient. The Sobel edge detector uses a pair of $3 \times 3$ convolution masks as shown below. The first mask estimates the gradient in the x-direction and the other estimates the gradient in the y-direction.

$$X \text{ gradient mask: } \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad Y \text{ gradient mask: } \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

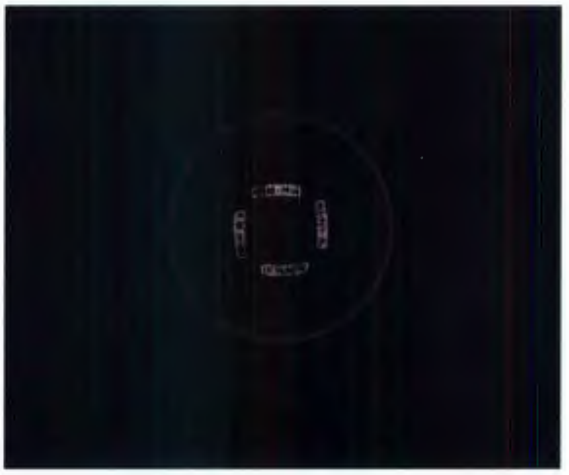
$$\text{Magnitude: } M = \sqrt{X^2 + Y^2} \tag{6.6}$$

$$\text{Direction: } \phi = TAN^{-1}\left[\frac{Y}{X}\right] \tag{6.7}$$

In this application, Sobel edge detection is only applied to the area of the image that contains the image of the sphere. Figure 6.10 shows the "gradient image" obtained by applying Sobel edge detection. It can be seen that there is a sharp change in the image intensity in the vicinity of the edge of the sphere.

(a)

(b)



(c)

Figure 6.10: Gradient images using Sobel method
(a) X gradient image (b)Y gradient image (c) Image showing magnitude

7) Determine the location of the edges to the nearest pixel: As mentioned above, edges are the areas with an elevated spatial derivative. From the estimated centroid, search in the radial direction for the maximum value of the pixel in the gradient (magnitude) image to determine the location of the edges to the nearest pixel. For greater efficiency, the

search is limited between *radius_minimum* and *radius_maximum* defined above for angle $\theta$ equals 0 to 180. The value of $\theta$ is incremented by a small value of 0.01 for each search. The coordinates of the pixels between *radius_minimum* and *radius_maximum* are estimated using the general equation for line-circle intersection. This method of searching from *radius_minimum* to *radius_maximum* for an angle $\theta$ from 0 to 180 will give two opposite points in the 2-D image where the circle and the line intersect.

The general equation of a line ($y = mx+c$) substituted in the equation of a circle of radius, $R$, and center coordinates ($\overline{X},\overline{Y}$) can be expressed as:

$$\left(x-\overline{X}\right)^2+\left((mx+c)-\overline{Y}\right)^2=R^2 \tag{6.8}$$

which can be simplified to the form of a quadratic equation:

$$Ax^2 + Bx + C = 0 \tag{6.9}$$

The two solutions for '$x$' ($x_1$, $x_2$) can be obtained using the formula:

$$x=\frac{-B\pm\sqrt{B^2-4AC}}{2A} \tag{6.10}$$

The two solutions for '$y$' ($y_1$,$y_2$) can be obtained by substituting the value of '$x$' into the general line equation:

$$y_1 = mx_1 + c \tag{6.11}$$

$$y_2 = mx_2 + c \tag{6.12}$$

These '$x$' and '$y$' coordinates will give the points in the gradient image where the line intersects the circle in the area between *radius_minimum* and *radius_maximum*. From these points obtain the coordinates of the maximum pixel for each search, i.e. peaks which will give the location of the edges to the nearest pixel. The location of these peaks will be refined later through subpixel interpolation for subpixel accuracy. Figure 6.11 shows the thresholded image obtained by leaving the edges. Here, all the peaks are given a pixel value '1', thereby representing a binary image.

Figure 6.11: Binary image of the calibration sphere showing peaks

8) Subpixel interpolation technique for subpixel accuracy: The peaks obtained by the above process are stored in a matrix to apply subpixel interpolation and thereby to

determine the edges in the image with subpixel accuracy. Subpixel interpolation is based on "fitting" a continuous interpolation function to the discrete gradient values (Krouglicof, [Online]). For this consider a Gaussian filter, i.e. convolution with a Gaussian kernel, $h(x,y)$ expressed as:

$$s(i,j) = f(i,j) \times h(x,y) \tag{6.13}$$

where the kernel or mask $h(x,y) = \frac{1}{2\Pi\sigma^2} e^{\frac{1}{2\sigma^2}(x^2+y^2)}$

The Gaussian filter can soften the edges depending on the value of $\sigma$. One technique for subpixel interpolation consists of determining the position of the mask that maximizes the value of the convolution evaluated using the gradient image and a Gaussian kernel. The steps involved in performing the subpixel interpolation can be summarized as follows:

- Perform the convolution of the gradient image with a Gaussian kernel. This can be expressed as:

$$s(x,y) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} g(i+m, j+n) \cdot e^{-\frac{1}{2}\left[(x-m)^2+(y-n)^2\right]} \tag{6.14}$$

- The position of the mask that maximizes the value of the convolution is given by the coordinates $(x,y)$. In order to determine these coordinates, the derivative of

$s(x,y)$ must be determined with respect to each of the unknown parameters and set equal to zero. This yields two nonlinear equations and two unknowns; i.e.

$$F_1 = 0 = \frac{\delta s(x,y)}{\delta x} = G \cdot (-1) \cdot (x - m) \tag{6.15}$$

$$F_2 = 0 = \frac{\delta s(x,y)}{\delta y} = G \cdot (-1) \cdot (y - n) \tag{6.16}$$

where $G = \sum_{m=-2}^{2} \sum_{n=-2}^{2} g(i+m, j+n) \cdot e^{-\frac{1}{2}\left[(x-m)^2 + (y-n)^2\right]}$

- Since these equations are nonlinear in nature, a numerical technique like Newton's method can be applied to solve these set of nonlinear equations. In general, Newton's method for solving nonlinear system of equations operates in an iterative manner, i.e. a vector '$\beta$' is found which satisfies the equation:

$$\overline{\beta} = -F(\overline{\theta}_0) J(\overline{\theta}_0)^{-1} \tag{6.17}$$

where $F(\overline{\theta}_0)$ represents the vector of function values (in this case it consists of two equations forming two rows and one column) and $J(\overline{\theta}_0)$ represents the Jacobian matrix which is defined as the matrix of partial derivatives of each of the *n* equations in the function vector evaluated with respect to each of the unknown parameters (in this case, it is a matrix containing two rows and two columns). Both $F(\overline{\theta}_0)$ and $J(\overline{\theta}_0)$ are evaluated using the most recent approximation to the

unknown parameters, $\overline{\theta}_0$. The new approximations to the unknown parameters are then computed by adding $\overline{\beta}$ and $\overline{\theta}_0$ as follows:

$$\overline{\theta} = \overline{\theta}_0 + \overline{\beta} \tag{6.18}$$

This step proceeds in an iterative fashion and continues until the Euclidean norm of the vector $\overline{\beta}$ is less than a certain threshold $\xi$ (i.e. specified accuracy criterion is attained) or until the maximum number of iterations is exceeded (i.e. the algorithm fails) (See Appendix D for a detailed explanation of the algorithm). Here, the *i* and *j* in the function vector (Equation 6.15 and Equation 6.16) can be substituted with the stored $y$ and $x$ coordinates of the peaks in the gradient magnitude image respectively.

- In order to apply Newton's method, the Jacobian matrix is evaluated by taking the partial derivative of each of the two equations (Equation 6.19 and Equation 6.20) with respect to each of the unknown parameters namely $x$ and $y$. The elements of the Jacobian matrix are given by:

$$\frac{\delta F_1}{\delta x} = G \cdot (x-m)^2 - G$$
$$\frac{\delta F_1}{\delta y} = G \cdot (x-m) \cdot (y-n) \tag{6.19}$$

$$\frac{\delta F_2}{\delta y} = G \cdot (y-n)^2 - G$$
$$\frac{\delta F_2}{\delta x} = G \cdot (x-m) \cdot (y-n)$$

(6.20)

These x and y parameters estimated through subpixel interpolation using the Newton's method represent the location of edges with subpixel accuracy (Figure 6.12). Blue circle points represent the location of edges refined through subpixel interpolation.

Figure 6.12: Edge points of the sphere refined through subpixel interpolation
(a) Location of edges with subpixel accuracy (b) Closer view of the location of edges represented by blue circular points

9) Nonlinear least square analysis: Once the location of edges with subpixel accuracy is obtained, the final procedure is to best-fit a circle with these set of data points using the concept of nonlinear-least square analysis. Nonlinear-least squares analysis refers to the problem of fitting a nonlinear model to a set of experimental data points (Krouglicof,

[Online]). It operates in an iterative manner, i.e. a vector $\beta$ is found which satisfies the equation:

$$(\overline{\beta}) = -\left(J(\overline{\theta}_0)^T J(\overline{\theta}_0)\right)^{-1} J(\overline{\theta}_0)^T F(\overline{\theta}_0) \tag{6.21}$$

where $F(\overline{\theta}_0)$ represents the vector of function values and $J(\overline{\theta}_0)$ represents the Jacobian matrix for the system of nonlinear equations under study. Both are evaluated using the most recent approximation to the unknown parameters, $\overline{\theta}_0$. The new approximations to the unknown parameters are then computed by adding $\overline{\beta}$ and $\overline{\theta}_0$ as follows:

$$\overline{\theta} = \overline{\theta}_0 + \overline{\beta} \tag{6.22}$$

This step proceeds in an iterative fashion and continues until the Euclidean norm of the vector $\overline{\beta}$ is less than a certain threshold $\xi$ (i.e. specified accuracy criterion is attained) or until the maximum number of iterations is exceeded (i.e. the algorithm fails) (Refer to Appendix E for a detailed explanation of nonlinear least-square analysis). Using this method of nonlinear least-square analysis, the unknown parameters, namely the center coordinates of the circle and the radius are estimated with subpixel accuracy and a circle is fitted for the data points. Figure 6.13 represents the circle fitted on the image of the sphere (highlighted in red) with subpixel accuracy for the set of data points (highlighted in yellow).


Figure 6.13: Best-fit circle for the set of edge points

### 6.2.3 Lining up of Spheres

The first step in the proposed technique was to line up the spheres kept at two different levels to determine the image center of the camera approximately. It can be executed by taking the image of the sphere kept at a cylindrical base of smaller height and then taking the image of the same sphere kept at the same location on the same base of higher height. Two images captured for a specific location and height can be processed using image processing algorithms as explained in the Section 6.2.2 to calculate the center of the circle. This process can be repeated for several locations in the world frame (i.e. where the spheres are kept) within the field of view of the camera until the center of the both

spheres in the image coincides at a common point. When the center of both spheres coincides it can be recognized that the both spheres were lined up along the optical axis of the camera, which will give an approximate image center of the camera initially. These image coordinates can be used later to obtain a mathematical model for calculating the exact image center and thereby the focal length of the camera, by considering the assumption that the plane of the table on which the spheres are kept and the image plane are not exactly parallel to each other. Figure 6.14 shows the schematic representation of the problem.

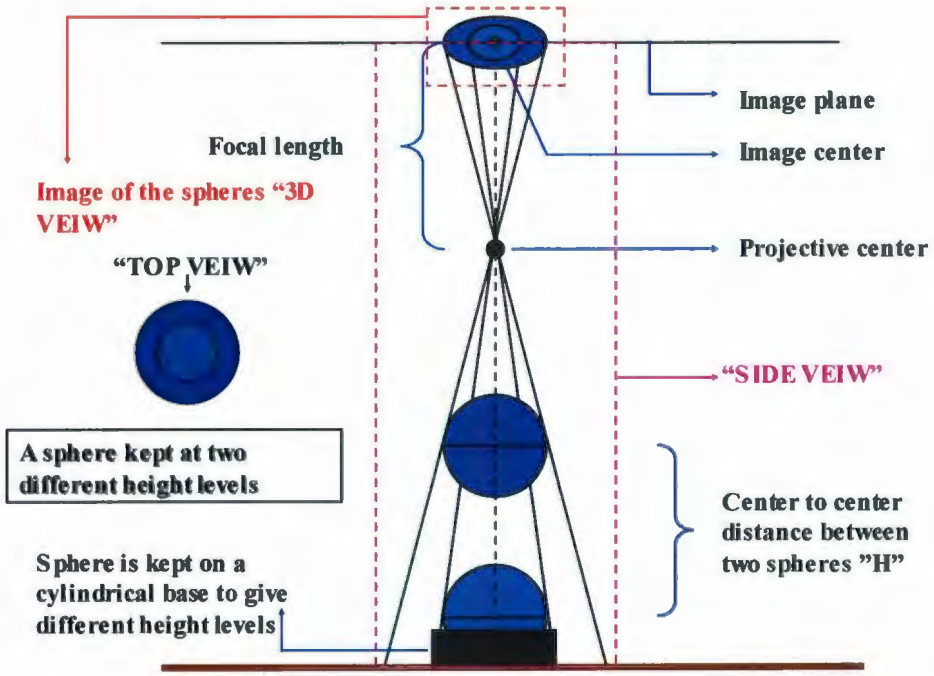



Figure 6.14: Schematic representation of the sphere line-up problem

## 6.2.4 Estimation of the Image Center and Focal Length of the Camera

### 6.2.4.1 Geometrical and Mathematical Models

In the real world, the plane of the object and the image plane may not be exactly parallel. Hence, once the spheres are lined up, the results obtained will be a rough approximation of the image center of the camera ($X_0$, $Y_0$). If both planes are parallel to each other, the obtained image center will coincide with the exact image center of the camera [($X_0$, $Y_0$) = ($U_0$, $V_0$)] (see Figure 6.15). In this figure it can be found that there are two displacements at a distance '*W*' taken towards the positive and negative side from the centered spheres. Each of the coordinates in the image plane represents the center of the circle, i.e. the image of the target sphere, and can be determined by processing the image of each sphere using the procedures as explained in Section 6.2.2.

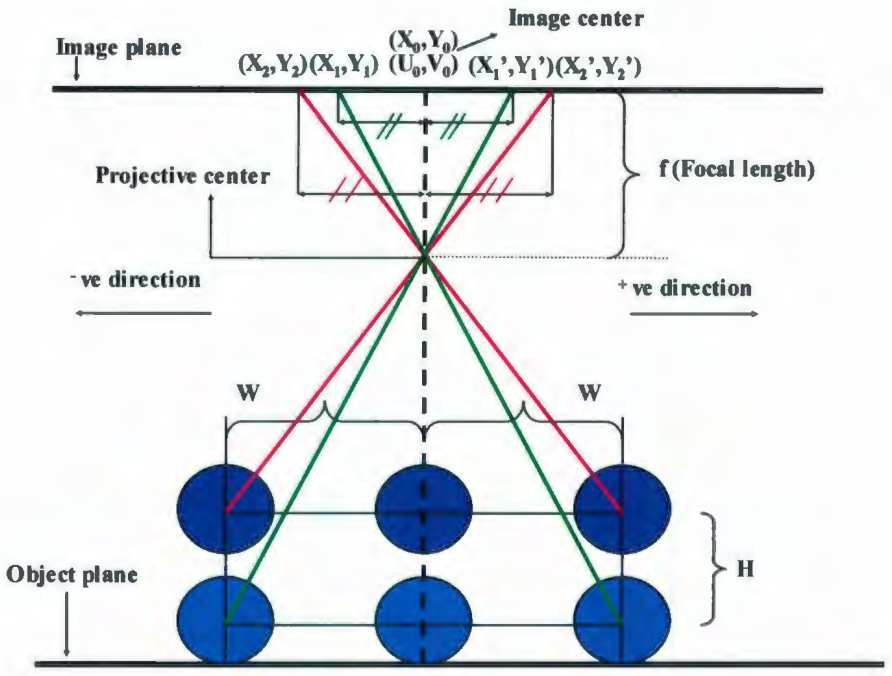



Figure 6.15: Geometrical model when the image plane and table are parallel

If the assumption is that both planes are parallel, then the image of these spheres after displacements will lie at opposite sides of the image center with an equal distance from the center. Hence, by identifying the image locations with respect to the center of the image plane, the slope or the direction of tilt of the image plane can be found.

Suppose that the two planes are not parallel. In order to quantify this assumption, another geometrical model is considered as shown in Figure 6.16. In the figure it is considered that the image plane is not parallel to the object plane and is tilted by a very small angle '$\alpha$'. From the accuracy point of view, this angle is not considered to be negligible in the camera calibration procedure. It can be noticed that in this situation once the spheres are lined up there is a small variation in the image center from the original which will depend on the angle '$\alpha$'. When there is a slope in the image plane, the image

of the spheres after displacements in the positive and negative direction will not lie at equal distances on opposite sides of the image center as mentioned before. To be more efficient in generating a mathematical model, only the x-coordinates of the image points are denoted which is obtained by moving the spheres to both the positive and negative side by a distance '*W*' along the X direction of the table in the test rig. In this case, the image plane is assumed to be as one-dimensional, where each x-coordinate of the image points obtained in the image plane lie in a line.

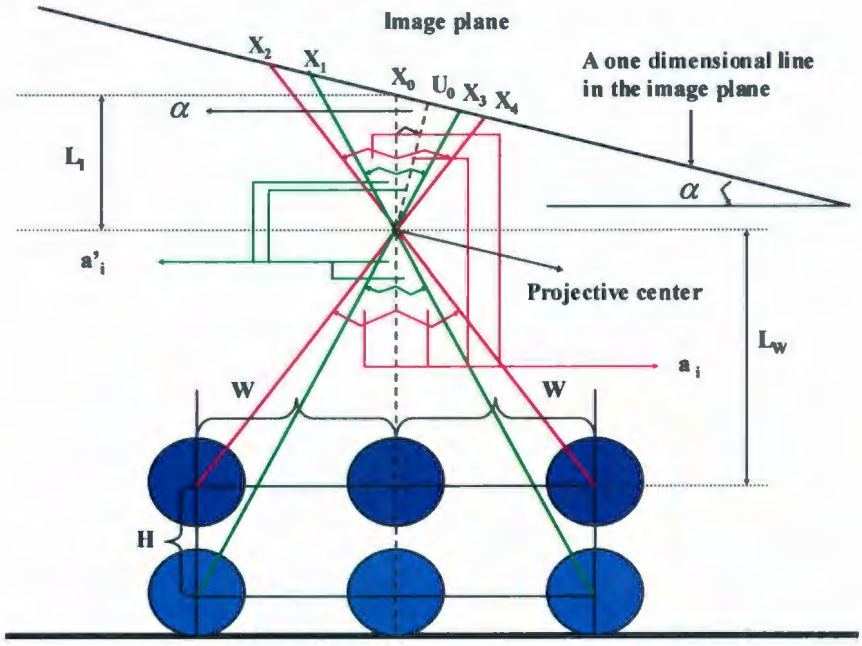



Figure 6.16: Geometrical model when the image plane and table are not parallel

The distance from the projective center to the point '$X_0$' is taken as '$L_1$', having length slightly higher than the focal length. The distance '$L_W$' represents the distance between the projective center and the center of the top sphere which is lined up. The distance between the centers of two spheres is represented by a height '*H*'. The angle made by the

axis of the top spheres is represented as '$a_t$' and those made by the bottom spheres is represented as '$a_t^{'}$'. The six equations generated from the model using the geometry is as follows:

$$L_l \tan(a_t + \alpha) = X_2 - U_0 \tag{6.23}$$

$$L_l \tan(a_t^{'} + \alpha) = X_1 - U_0 \tag{6.24}$$

$$L_l \tan(a_t - \alpha) = U_0 - X_4 \tag{6.25}$$

$$L_l \tan(a_t^{'} - \alpha) = U_0 - X_3 \tag{6.26}$$

$$\tan a_t = \frac{W}{L_W} \tag{6.27}$$

$$\tan a_t^{'} = \frac{W}{L_W + h} \tag{6.28}$$

The six unknown parameters in these equations are $L_l, L_W, \alpha, U_0, a_t, a_t^{'}$. Since the equations are nonlinear in nature, an analytical approach is not always feasible and hence a numerical technique like Newton's method is applied to solve these equations (See Appendix D for a detailed explanation of the Newton's method algorithm). The values of '*W*' and '*H*' are taken as 25 and 53.35mm respectively. Once $L_l$ and $\alpha$ are estimated using the above method, the focal length ($f$) can be calculated using the formula:

$$f = L_l \cos(\alpha) \tag{6.29}$$

The drawback of this method is that from the experiment it is assumed that only four significant digits will be obtaining for each of the image points. If that is the case, then the solutions for the unknown parameters will not be accurate.

Taking into consideration of the results obtained by Newton's method, another geometrical model is considered (Figure 6.17) where more displacements are provided to generate more image points by moving the spheres '$n$' times to one side (positive direction), with a distance '$W$' . For better understanding, displacements towards one direction are only shown in the figure. The direction shown is considered to be along the X-axis. The dotted lines towards right and left indicates that the displacements of the sphere are given continuously by a distance '$W$' each time in the object plane and corresponding image points are obtained in the image plane. For '$n$' displacements of the sphere '$2n$' image points are obtained.

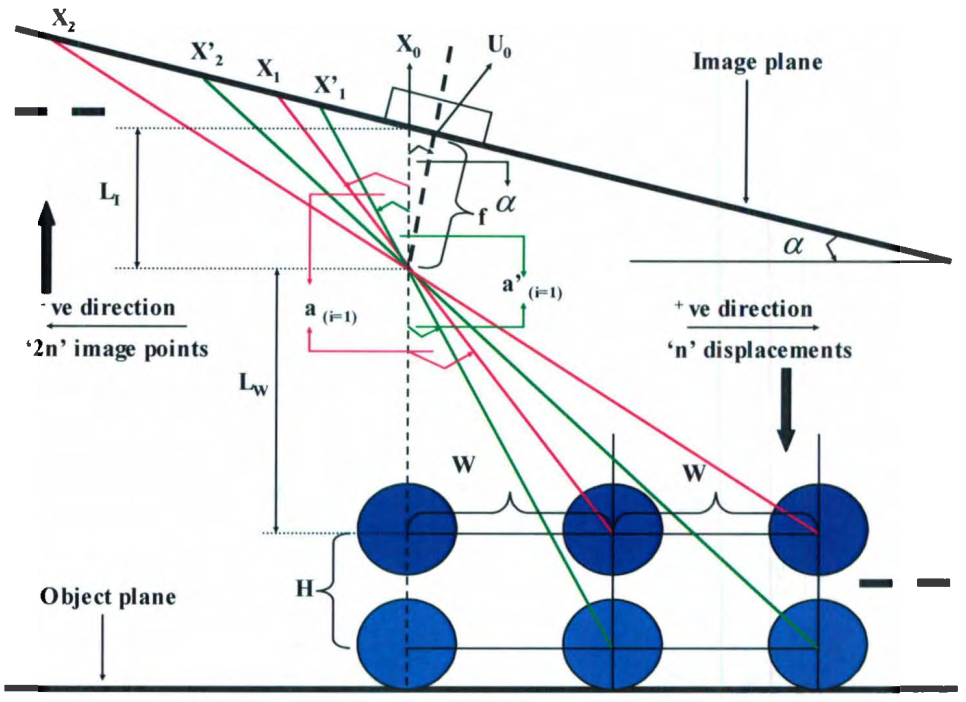



Figure 6.17: Geometrical model with 'n' displacements of the sphere to the right

From the geometry, four equations are generated for each point in the image plane for the displacements $i=1.....n$ along the X axis of the table in the test rig:

$$\tan a_i = \frac{W_i}{L_W} \tag{6.30}$$

$$\tan a_i^{'} = \frac{W_i}{L_W + h} \tag{6.31}$$

$$X_i = U_0 + L_1 \tan(a_i + \alpha) \tag{6.32}$$

$$X_i^{'} = U_0 + L_1 \tan(a_i^{'} + \alpha) \tag{6.33}$$

The points $X_i$ and $X_i$' represent the x-coordinates of the 2-D image points obtained in the image plane. The unknown parameters in the above equations are $L_l, L_w, \alpha, U_0, a_i, a_i^{'}$. For displacements, $W_{i=1..n}$, there will be angles $a_i$ and $a_i^{'}$ for $i=1...n$. Hence for each displacements, additional $a_i$ and $a_i^{'}$ are added to the list of unknown parameters along with the other four parameters namely $L_l, L_w, \alpha, U_0$, which are common in all displacements. Therefore to summarize, if there are '$n$' displacements, there are '$4n$' equations and '$2n+4$' unknowns. If the displacements are taken towards the negative direction, the same four equations are sufficient to solve these unknown parameters. Hence, to fit a nonlinear model to this set of experimental data points (displacements), a nonlinear least square analysis problem has been used. Refer Appendix E to see the description of the algorithm of nonlinear least square analysis. Similarly, these four equations can be used to solve the unknown parameters when the table is moved along the Y-axis (i.e. perpendicular to the above movement of the table) where the points $X_i$ and $X_i$' will be substituted with the obtained $Y_i$ and $Y_i$' points for both spheres and $U_0$ will be substituted with $V_0$. Equations (6.32) and (6.33) will be then be changed to:

$$Y_i = V_0 + L_l \tan(a_i + \alpha) \tag{6.34}$$

$$Y_i^{'} = V_0 + L_l \tan(a_i^{'} + \alpha) \tag{6.35}$$

where $Y_i$ and $Y_i$' represents the y-coordinates of the image points for $i=1..n$ displacements and $V_0$ represents the y-coordinate of the image center.

Once $L_1$ and $\alpha$ are estimated, the focal length ($f$) can be calculated using Equation (6.29). It is recognized from the simulation that, the more the displacements are added the more accurate the parameter values are obtained. Therefore, using this concept, the results are verified by considering 50 displacements to the positive direction and negative direction, which gave comparatively good results for the unknown parameters in the simulation. This was then implemented experimentally which will be described in detail in the next section.

### 6.2.4.2 Experiments

The results obtained using simulation were verified by undergoing some real-time experiments. Figure 6.18 shows the complete experimental set up for the proposed technique.

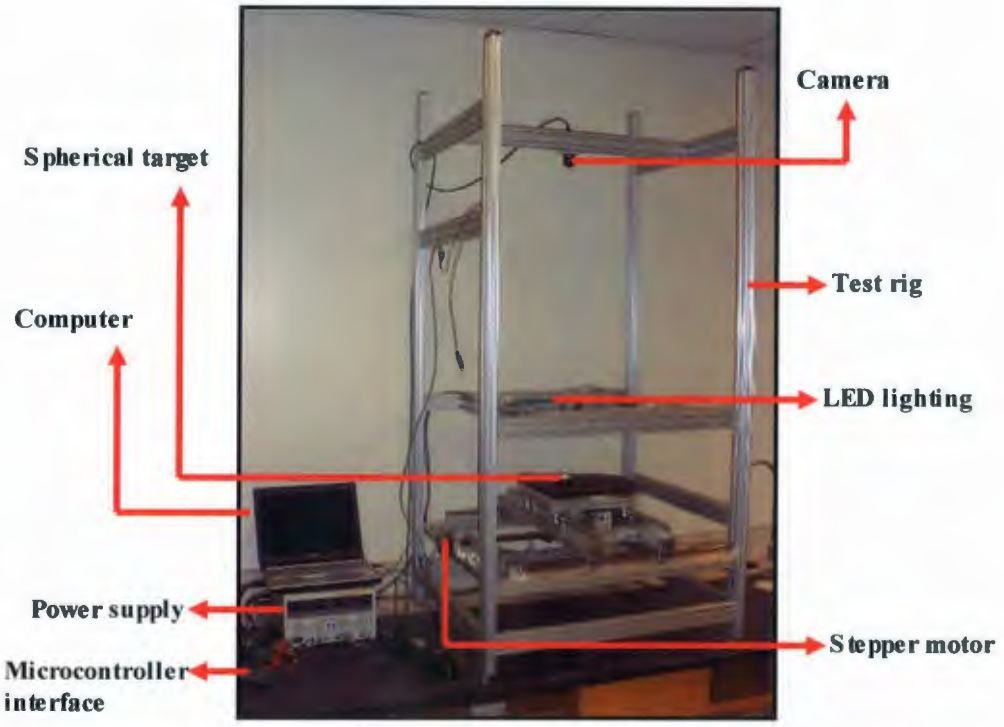



Figure 6.18: Complete experimental setup

As described before, the first step in the proposed technique is to line up the spheres kept at two different levels to determine the image center of the camera approximately. It was executed by taking the image of the sphere kept at a cylindrical base of smaller height and then taking the image of the same sphere kept at the same location on the same base of higher height. The experimental set up used for this case was exactly in the same way as shown in the Figure 6.18. The microcontroller interfacing circuit along with the stepper motor driven ball screws controlled the motion of the table in the X (horizontal) and Y (vertical) directions. A Graphical User Interface (GUI) which includes a data acquisition software along with the Lumenera MATLAB driver were implemented to

capture the images of the spheres. The GUI includes features that automate the process of motion of the X-Y table in accordance with the user's information and capturing of images of the target. Two images obtained on the image plane for a specific location were processed using image processing algorithms (as explained in Section 6.2.2) to calculate the center of the circle. The center coordinates thus obtained were examined each time to acquire the information regarding the direction to which the motors have to get rotated in the next step. This process was repeated for several locations in the object plane (i.e. where the spheres were kept) within the field of view of the camera until the center of both spheres in the image coincides at a common point. When the center of both spheres coincides it was recognized that both spheres were lined up along the optical axis of the camera, which gave an approximate image center of the camera initially. The spheres were lined up when their image coordinates were approximately (712.6894, 429.8865) and (712.5424, 429.4762) pixels for the bottom and top sphere respectively. Once the spheres were lined up, to check for correctness, a few trials were taken at the same location in order to obtain the repeatable results. Table 6.1 shows the trial runs performed. $X_0$ and $Y_0$ represent the x and y image coordinates of the image center and R represents the radius of the circle obtained on the image plane. The results were almost similar. Small variations are due to noise in the camera. The underlying fact here is that multiple images will reduce the noise which forms the next step of the experiment.

**Table 6.1: Trial runs for lining up the spheres**

| | Bottom Sphere | | | Top Sphere | | |
|---|---|---|---|---|---|---|
| Trial | $X_0$ | $Y_0$ | R | $X_0$ | $Y_0$ | R |
| 1 | 712.6939 | 429.8497 | 109.1571 | 712.5473 | 429.4547 | 116.0257 |
| 2 | 712.6627 | 429.8213 | 109.1519 | 713.0042 | 429.1541 | 116.0301 |
| 3 | 712.6449 | 429.8481 | 109.1602 | 713.0034 | 429.1779 | 116.0268 |

Once the spheres were lined up the next step undertaken was to move the spherical target towards the positive and negative directions in X and Y. The sphere was placed at two different heights as before. As mentioned earlier, 50 displacements were taken in the positive and negative X and Y axis directions for both heights from the centered position. Images were captured by camera for each displacement resulting in a total of 400 images. These images were stored in a particular directory in the computer for further processing. These images were processed using a series of image processing algorithms plus several other techniques as explained in Section 6.2.2 for estimating the center coordinates of each image. The center coordinates in pixels, thus obtained with subpixel accuracy through subpixel interpolation were multiplied with the pixel size of the camera represents the $X_i$ and $X_i'$ as well as $Y_i$ and $Y_i'$ points in mm. These 400 points were then used to calculate the unknown parameters namely $L_1, L_W, \alpha, U_0, a_i, a_i'$ for $i = 1$ *to* $n$ displacements in X-axis ($L_1, L_W, \alpha, V_0, a_i, a_i'$ for $i = 1$ *to* $n$ displacements in Y-axis) using nonlinear least square analysis (refer to Appendix E to see the description of the algorithm of nonlinear least square analysis). The pixel size of the camera in X and Y was taken as $P_x$ = 0.00465mm and $P_y$ = 0.00465mm. In this case n = 101, i.e. 50 displacements to the left and 50 displacements to the right plus the center location in each

axis direction. For 101 displacements, the value of 'W' was incremented by 0.5mm which makes the target to be within the field of view of the camera. The distance, 'H', i.e. the center to center distance between the bottom sphere and top sphere was taken as 53.35mm. For the nonlinear least squares analysis problem, initial approximations of the unknown parameters have to be set. Hence the initial parameter values of focal length, 'F' and '$\alpha$' were taken as 38mm and 0 degrees from which the initial estimate of $L_I$ was calculated by using the formula obtained from geometry, $L_I = \frac{F}{\cos(\alpha)}$, i.e. $L_I$ = 38mm. The total distance from the projective center (also known as focal point) to the bottom sphere was approximately taken as 900mm. Hence the initial estimate for $L_W$ was taken as $L_W$ = 900-H. The initial value for $U_0$ and $V_0$ were taken as 3.31359mm and 2mm, respectively which were obtained from the $X_0$ and $Y_0$ value when the spheres were lined up. The initial values for $a_i, a_i'$ were obtained by using Equations (6.30) and (6.31), respectively. The unknown parameters were then estimated separately for the displacements in X and Y axis of both spheres by applying the algorithm of nonlinear least squares analysis.

### 6.2.4.3 Results and Discussions

The results of the above experiment indicate that the algorithm of nonlinear least squares analysis (Appendix E) works perfectly in calculating the unknown parameters namely $L_I, L_W, \alpha, U_0(V_0), a_i, a_i'$. As mentioned in Section 6.2.4.1 these unknown

parameters were estimated separately for the displacements in X and Y axis of both spheres. The following results were obtained from the technique undertaken:

*Displacements of bottom and top sphere in the X-axis direction:*

The problem of nonlinear least square analysis is to fit a nonlinear model to a set of experimental data points (n=101). It was noted that the algorithm had converged within four iterations. Since there are 206 unknown parameters (i.e. m=2n+4), the unknown parameter vector to be estimated consists of a size of 206 rows and 1 column. The obtained values for $a_i$ range from -0.0288 to +0.0288 and those for $a_i^{'}$ range from -0.0271 to +0.0272. From geometry, it is clear that the angle $a_i$ (i.e. the angle made by the top sphere) should be greater than $a_i^{'}$. Hence, the values obtained for the angles from the experimental study can be recognized as valid results. Similarly, the obtained values for the other unknown parameters are: $L_I$ = 36.6599mm, $L_W$ = 867.5287mm, $\alpha$ = 0.0065 radians (i.e. 0.3724 degrees), $U_0$ = 3.0761mm (i.e. 661.53 pixels). The value of $\alpha$ indicates that the object plane and image plane are not parallel to each other and that there is a small tilt in the image plane. From $L_I$, the focal length obtained was 36.6599mm which matches with the focal length obtained using the Heikkila's method of camera calibration (refer to Table 6.2). The value of 867.5287mm obtained for $L_W$ also seems to be correct for the distance from the projective center to the top sphere. The camera used for the experiment was Lumenera (Lm135) which had a resolution of 1392*1040 pixels. In that case, the value of the x-coordinate of the image center, i.e. $U_0$ = 3.0761mm (661.53 pixels) is also observed to be a valid result.

The accuracy of the data acquisition process plays a vital role when evaluating the final results of the camera calibration technique. For validation and estimation of error in the image points, the $X_i$ and $X_i$' points were calculated using the obtained unknown parameters as initial estimates and using the mathematical model generated from geometry. Figure 6.19 shows a comparison of the $X_i$ and $X_i$' (in the MATLAB figure, $X_i$' is denoted by the notation $X_i$") image points of top and bottom spheres, respectively. The plotted blue lines represent the image locations obtained by conducting the experiment and the red lines represent those values obtained by applying the analytical model developed from geometry. It can be seen that both values of image points follow similarly in each displacement. The percentage error in $X_i$ points falls within a range of 0.04% to +0.13% and that in $X_i$' points falls within a range of -0.01% to +0.06%, which are negligibly small and acceptable (Figure 6.20). These results show the experimental set up corresponds with the mathematical model developed from the geometry in calculating the 2-D image locations in the image plane.

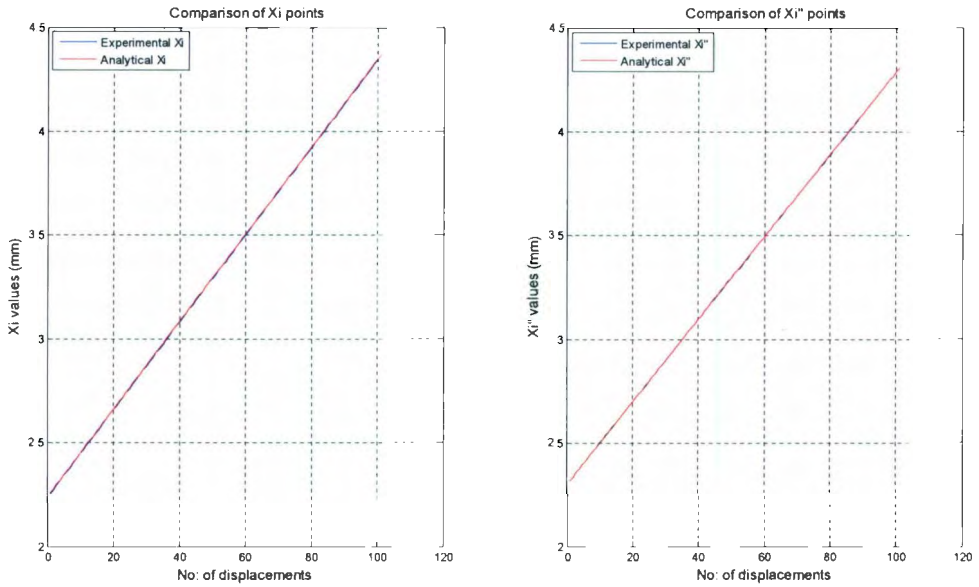



Figure 6.19: Comparison of image points (X-axis displacements)

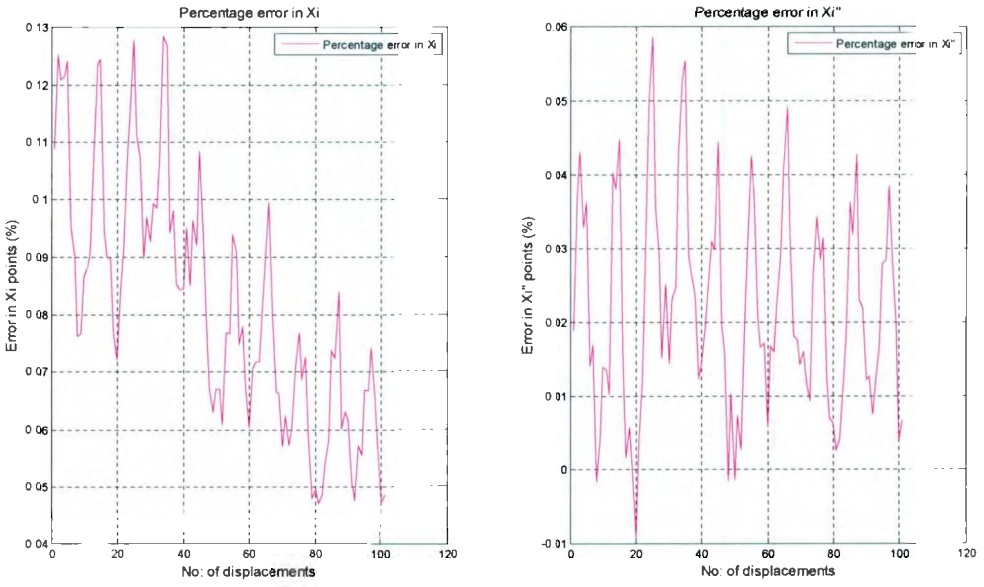



Figure 6.20: Percentage error in the image points (X-axis displacements)

*Displacements of bottom and top sphere in the Y-axis direction:*

In this case, both the top and bottom spheres are moved in the positive and negative directions from the image center, but the displacements are given in the Y-axis. The results obtained were similar to the X- direction. Here, the algorithm of nonlinear least squares analysis had converged within five iterations. The values obtained for each of the unknown parameters are as follows: The values range from -0.0288 to +0.0288 for $a_i$ and from -0.0272 to +0.0271 for $a_i'$. The values estimated for $L_I$, $L_W$, $\alpha$, and $V_0$ were 36.6989mm, 867.8981mm, -0.0183 radians (i.e. 1.05 degrees) and 2.6720mm (i.e. 574.62 pixels), respectively. The focal length calculated from $L_I$ corresponds to 36.6989mm. This value matches with that obtained above for the X displacements. Similar results were observed corresponding to $L_W$. A negative value of $\alpha$ indicates that the image plane is tilted by a small angle in the opposite direction with respect to the value of $\alpha$ during the X-axis displacements.

As mentioned earlier, for validation and estimation of error in the image points, the $Y_i$ and $Y_i$' points were calculated by using the obtained unknown parameters as initial estimates and by using the mathematical model generated from the geometry. The Figure 6.21 shows a comparison of the $Y_i$ and $Y_i$' (in the MATLAB figure, $Y_i$' is denoted by the notation $Y_i$") image points of top and bottom spheres respectively. The plotted blue lines represent the image locations obtained by conducting the experiment and the red lines represent those locations obtained by applying the analytical model developed from the geometry. It can be seen that both values of image points follows are similar for each displacement. The percentage error in $Y_i$ points falls within a range of -0.15% to +0.1%

and that in $Y_i$' points falls within a range of +0.05% to +0.35 %, which are comparatively negligible and acceptable (Figure 6.22). Hence, this also exemplifies that the experimental set up used corresponds with the mathematical model developed from the geometry in calculating the 2-D image locations in the image plane.

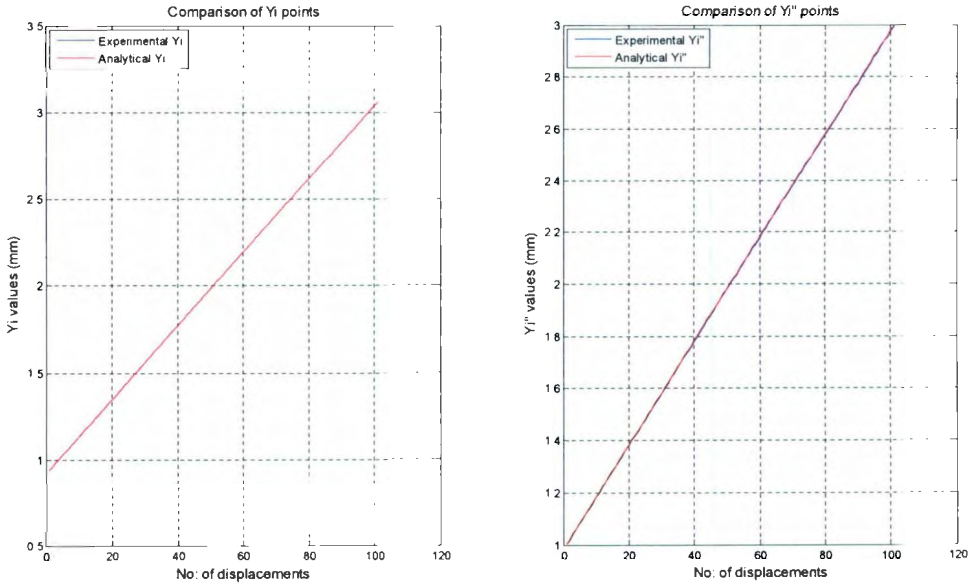



Figure 6.21: Comparison of image points (Y-axis displacements)

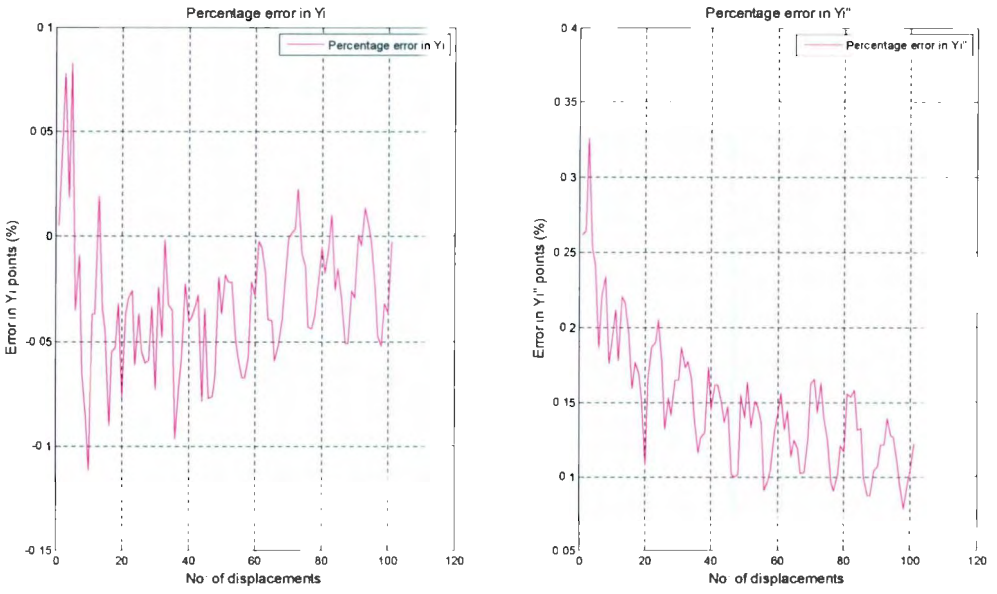



Figure 6.22: Percentage error in the image points (Y-axis displacements)

Using this new approach, the image center of the camera was calculated as (661.53, 574.62) pixels and focal length was 36.6mm. The results obtained by this technique were compared with the results obtained using Heikkila's camera calibration algorithm in Table 6.2.

**Table 6.2: Comparison of the estimated parameters**

| Parameters | X-axis displacements | Y-axis displacements | Heikkila's Method* |
|---|---|---|---|
| $f$ (mm) | 36.6599 | - | 36.6597 |
| $f$ (mm) | - | 36.6989 | 36.7257 |
| $U_0$ (pixels) | 661.53 | - | 756.27 |
| $V_0$ (pixels) | - | 574.62 | 630.38 |
| $\alpha_x$ (degrees) | 0.3724 | - | 0.423 |
| $\alpha_y$ (degrees) | - | 1.05 | 1.5 |
| $L_W$ (mm) | 867.52 | 867.89 | 859.63 |
| $s$ | 1.001 | | 1.0018 |

* Taken from "T. Rahman, 2009: Highly Accurate Vision Based Surface Reconstruction System (Thesis Report), pp. 139".

In the table, for differentiating the tilted angle, '$\alpha_x$' and '$\alpha_y$' are used for X and Y displacements, respectively. The intrinsic parameter 's' is calculated by taking the ratio between the focal length obtained in the Y-axis displacements to X-axis displacements. The parameter values mentioned in the last column of the table is acquired by using the camera calibration algorithm proposed by Heikkila with 4800 target points. These are used as a reference for the comparison of this new technique. It can be seen from the table that the focal length obtained using this proposed technique is similar to that obtained by Heikkila's method. Hence, this technique can be considered to be more efficient in calculating the focal length precisely with only a less number of target images (in this case: 202 images) compared with the Heikkila's algorithm using 4800 target

images. Similarly the calculated scale factors are comparable from the two methods. But in the case of image center, there is a difference of around 95 pixels for $U_0$, and around 55 pixels for $V_0$ which cannot be quantified unless the true value of the image center is known. This suggests further improvements need to be done in the proposed new technique for making a conclusion regarding the accuracy of the image center. Regarding $\alpha$, there is only a slight variation in the angle calculated from both techniques. Similar with the case of $L_W$, where there is a small difference of around 8mm in length.

To summarize, using the proposed technique the intrinsic parameters like image center and focal length are computed separately without taking into consideration lens distortions of the camera. This makes the technique simple and efficient in estimating the camera parameters without involving the incorporation of any complex nonlinear lens distortion models. Also, the time required to undertake the technique is comparatively less and hence is suitable for applications that require high speed calibration. The observations from the experimental study show that this new technique for separately finding these parameters of the camera has the potential to improve the calibration accuracy especially in the case of calculating focal length whereas further improvements are recommended to get a conclusion on estimating the image center precisely.

# Chapter 7

# Conclusions and Recommendations

In the field of machine vision, camera calibration refers to the experimental determination of a set of parameters which describe the image formation process for a given analytical model of the machine vision system. In 3-D machine vision, it is necessary to know the relationship between the 3-D object coordinates and the 2-D image coordinates. This transformation is determined in Geometric Camera Calibration (GCC) by experimentally determining the unknown parameters of the camera model. It introduces the analytical tools necessary to establish the quantitative constraints between the image measurements and the position and orientation of geometric figures measured in some arbitrary coordinate system. Once a camera has been calibrated, it is possible to associate with an image point a well defined ray which passes through this point and the camera's optical center as well as to perform three dimensional measurements from a

digitized picture. An accurate, reliable calibration procedure is essential for most industrial machine vision applications including areas like mechanical metrology, robot assembly, reverse engineering, image fusion, tracking, etc. An optimal calibration technique should produce unbiased and minimum variance estimates of the camera parameters. In practice, this is quite difficult to achieve due to various error sources affecting the imaging process. The precision of the parameter estimates is an important aspect in all camera calibration techniques. Hence, geometric camera calibration followed by the estimation of parameters precisely is not a trivial task in any vision based applications. Therefore, this study mainly focused on the role of "accuracy" in geometric camera calibration by identifying a few predominant factors contributing to error. Also, a new approach towards camera calibration is proposed in which image center and focal length is calculated independently of the lens distortion.

## 7.1 Summary of Results and Conclusions

The following paragraphs explain the research work undertaken and the conclusions that can be drawn from the work:

- The first problem addressed in this thesis was to determine whether the algorithm reported in Heikkila (2000), works properly in estimating the camera parameters assuming ideal data (i.e. no measurement errors). This was verified from the calibration results obtained, which illustrated that the algorithm works perfectly in determining the camera's extrinsic and intrinsic parameters.

- As a second step, an accuracy assessment was performed on Heikkila's method. The problem undertaken was to gain an understanding of how much accuracy is needed on the target points for an optimal calibration which addresses the following fundamental question: Assuming a certain tolerance or uncertainty in the calibration target, what is the expected error with respect to the measured camera parameters and what is the impact on the final 3-D machine vision application? The relative accuracy obtained for each parameter when incorporating an uncertainty of 0.1mm on each of the target points has been reported as an outcome of the study. From the histogram plots and the accuracy analysis, it was observed that the statistical dispersion of distortion parameters is high in comparison to the other intrinsic parameters and that the order of magnitude of the variation is close to the actual value of the parameter. An important conclusion that can be drawn from this is that a very high accuracy is required on the target points in order to accurately determine the distortion parameters. Also, it was noticed that a relatively small change in the rotation angle about the Z axis results in a relatively large change in the observed image. The accuracy assessment study clearly indicates that the dimensional accuracy of the calibration target plays an important role in geometric camera calibration and that a carefully machined "precision" calibration target is critical in applications demanding high accuracy. Therefore, the results illustrate that in a real world machine vision application it is extremely difficult to obtain accurate or

repeatable results even though current camera calibration techniques and algorithms are sound from a theoretical point of view.

- The next move in the research was to understand the problems associated with geometric camera calibration through the application of Design of Experiments (DOE). The camera calibration toolbox for MATLAB proposed by Heikkila is used for computer simulation. The primary (intrinsic) factors associated with camera calibration are emphasised in this experiment, namely, focal length, image centre, radial and tangential distortions. The response is taken as sum square of error caused by varying different combinations of the factors considered. A response surface methodology (RSM), namely a Central Composite Design (CCD), was carried out for this purpose. The objective of this study was to determine the significant factors leading to inaccurate results during the process of camera calibration. As an outcome of the study performed, significant findings are reported, including large interactions between certain factors that help explain the difficulties in obtaining requisite calibration results in many real-world machine vision applications where high accuracy is essential. The results indicate that due to the interactions between parameters the calibration algorithm will not necessarily converge to the minimum value when the data is subject to measurement error and the error cannot necessarily be attributed to a single camera parameter. Also, it was noted that the results from Design-Expert are acceptable with the results obtained from MATLAB calibration model, the interaction plots are reported for validation purposes. One of the most significant

findings is that certain camera parameters (e.g. radial distortion 2) can effectively be removed from the calibration model with no discernible loss in accuracy; a significant finding that goes against much of the published camera calibration literature.

- Based on these DOE results, the final focus of this thesis was to propose a new approach towards geometric camera calibration in which image center and focal length are calculated independently of the lens distortion. The idea behind this study was to differentiate certain intrinsic parameters like image center and focal length and compute them separately without incorporating any nonlinear lens distortion models. This forms a novel technique towards GCC involving mechatronics and the algorithms are based on state of the art image processing algorithms including subpixel interpolation techniques for subpixel accuracy and a few numerical techniques. The experimental results show that this new technique for finding the above mentioned camera parameters has the potential to improve the calibration accuracy especially in the case of calculating focal length whereas further improvements are recommended to get a conclusion on estimating the image center precisely. Also, this technique is simple, efficient and comparatively fast and hence is suitable for applications that require high speed calibration.

## 7.2 Contributions

The following is a summary of the main contributions of this work:

- An accuracy assessment on Heikkila's method is performed by adding an uncertainty to the target points and thereby to investigate the 3-D measurement errors.
- Novel application of Design of Experiments (DOE) as an aid in finding the problems associated with geometric camera calibration. As an outcome of the study performed, significant findings are reported, including large interactions between certain factors that help explain the difficulties in obtaining requisite calibration results in many real-world machine vision applications where high accuracy is essential. One of the most significant findings is that certain camera parameters (e.g. radial distortion 2) can effectively be removed from the calibration model with no discernible loss in accuracy; a significant finding that goes against much of the published camera calibration literature.
- Development of a novel technique to estimate certain intrinsic camera parameters like image center and focal length independently of the lens distortion based on the state of the art image processing algorithms including subpixel interpolation techniques for subpixel accuracy and a few numerical techniques. A GUI was developed and implemented for automating the process of capturing the images of the target. Also, this technique is simple, efficient and comparatively fast and hence is suitable for applications that require high speed calibration.

## 7.3 Recommendations

The following are the recommendations for future work:

- It was already discussed that the accuracy of data acquisition process plays a crucial role in geometric camera calibration when evaluating the final estimates of camera parameters. The precision of the final results of camera calibration depends on how accurately the camera model is used and how 3-D coordinate points and its corresponding 2-D image points are located. Therefore, the current work on accuracy and error analysis can be further extended to deal with the problem of error propagation of image coordinates on the camera calibration parameters resulting from the measurement errors of 3-D and 2-D coordinate points. An error propagation model can be derived for the camera model used in calibration procedure, and be used to study the impact of image coordinates error and how these errors are propagated in to the final results of camera calibration. Further, the impact of increasing the number of image points on the errors of image coordinates can also be investigated.
- In the study of design of experiments, a half CCD is used which requires less number of runs compared to full CCD. Hence, for further investigations a full CCD or a different type of RSM like BBD can be used for checking the correctness and accuracy of the model obtained. Also, the results from the DOE approach can be used for better understanding in vision based applications involving the GCC of cameras.

- In the new proposed technique of estimating certain camera parameters separately, the projection of the occluding contour of a sphere can be interpreted as an ellipse in the image. It is suspected that, this can be used for better efficiency and accuracy of the new proposed technique. This geometric interpretation needs the inclusion of conic fitting algorithms, namely ellipse fitting to extract the contour and thereby to obtain the control points from the images.
- Further extension to the new technique can be done by increasing the number of target points and thereby the corresponding image points. It is expected that this can reduce the error and improve the accuracy in the estimated camera parameters during calibration process.
- For a comparative study and to have further improvements, the calibration results from the proposed technique can be compared with the results obtained from the other calibration techniques.
- The various image processing algorithms and numerical techniques involved in the proposed new approach can be implemented for real-time applications.

# Bibliography


1. **"CAMO".**, [Online]. Available: http://www.camo.com/rt/Resources/design_of_experiment.html
2. **"Computer Vision".**, [Online]. Available: http://en.wikibooks.org/wiki/Robotics/Sensors/Computer_Vision
3. **"NIST/SEMATECH".**, e-Handbook of Statistical Methods. [Online]. Available: http://www.itl.nist.gov/div898/handbook/index.htm
4. **"NLREG".**, [Online]. Available: http://www.nlreg.com/index.htm
5. **"The Quality Portal".**, [Online]. Available: http://thequalityportal.com/q_know02.htm
6. **A. Gruen and H. A. Beyer.**, 2001: System calibration through self-calibration. *In 'Calibration and Orientation of Cameras in Computer Vision' Gruen and Huang (Eds.)*, Springer Series in Information Sciences 34, pp. 163-194
7. **A. Gruen and T. S. Huang**, (Eds.)., 2001: *Calibration and Orientation of Cameras in Computer Vision.* Springer Series in Information Sciences, Vol. 34
8. **A. Gruen and T.S. Huang.** (Eds.), 2001: *Calibration and Orientation of Cameras in Computer Vision.* Springer-Verlag, Heidelberg, Germany

9. **A. Ryberg, A-K Christiansson, K. Eriksson, and B. Lennartson.**, 2006: A new camera model and algorithms for higher accuracy and better convergence in vision-based pose calculations. *In Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation*, pp. 194-199

10. **B. Caprile and V. Torre.**, 1990: Using vanishing points for camera calibration. *International Journal of Computer Vision*, Vol. 4(2), pp. 127-140

11. **B. Green.,** 2002: "Edge Detection Tutorial"., [Online]. Available: http://www.pages.drexel.edu/~weg22/edge.html

12. **B. Guendouz, C. Eswaran, and S. V Muniandy.**, 2006: Error propagation and accurate calibration for camera model. *IEEE International Conference on Engineering of Intelligent Systems*, pp. 1-5

13. **B. S. Morse.,** 1998-2000: "Thresholding"., [Online]. Available: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/threshold.pdf

14. **B. Triggs.**, 1998: Auto calibration from planar scenes. *ECCV 98*, pp. 89-105

15. **B.G. Batchelor and P.F. Whelan**, 1997: *Intelligent Vision Systems for Industry*. Springer-Verlag

16. **C. C. Slama.**, editor., 1980: *Manual of Photogrammetry*. Fourth edition, American Society of Photogrammetry, Falls Church, Va.

17. **C. C. Wang.**, 1992: Extrinsic calibration of a vision sensor mounted on a robot, *IEEE International Journal of Robotics and Automation*, Vol. 8, pp. 161-175

18. **C. R. Hicks and K. V. J Turner.**, 1999: Fundamental concepts in the design of experiments. Oxford University Press, Inc.

19. **C. S. Fraser.**, 1997: Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 52, pp. 149-159

20. **C. S. Fraser.**, 2001: Photogrammetric camera component calibration. A review of analytical techniques. *In 'Calibration and Orientation of Cameras in Computer Vision', Gruen and Huang (Eds.)*, Springer Series in Information Sciences 34, pp. 95-121

21. **D. C. Brown.**, 1966: Decentering distortion of lenses. *Photogrammetric Engineering and Remote Sensing*, Vol. 32(3), pp. 444-462

22. **D. C. Brown.**, 1971: Close-range camera calibration. *PE &RS*, Vol. 37(8), pp. 855-866

23. **D. C. Brown.**, 1971: Close-range camera calibration. *Photogrammetric Engineering*, Vol. 37(8), pp. 855-866

24. **D. C. Brown.**, 1974: Evolution, application and potential of the bundle method of photogrammetric triangulation. *Proc. ISP Symp.*, pp. 69

25. **D. C. Montgomery.**, 2008: Design and Analysis of Experiments, 7th Edition, John Wiley and Sons Inc, NJ

26. **D.A. Forth and J. Ponce.**, 2003. *Computer Vision: A modern Approach*. Pearson Education Inc, Upper Saddle River, NJ 07458

27. **E. L. Hall, M. B. K. Tio, C. A. McPherson, and F. A. Sadjadi.**, 1982: Curved surface measurement and recognition for robot vision. *In Conference Record on IEEE Workshop Industrial Applications of Machine Vision*

28. **E. R. Davies.**, 2004: *Machine Vision : Theory, Algorithms, Practicalities*. Third edition, Morgan Kaufmann

29. **F. Remondino and C. Fraser.**, 2006: Digital camera calibration methods: Considerations and Comparisons. *IAPRS*, Vol. XXXVI, Part 5, Dresden 25-27

30. **G. Shivaram and G. Seetharaman.**, 1998: A New Technique for Finding the Optical Center of Cameras. *In Proceedings of the International Conference on Image Processing (ICIP'98)*, Vol. (2), pp. 167-171.

31. **G. Wei and S. Ma.**, 1993: A complete two-plane camera calibration method and experimental comparisons. *In Proceedings of the Fourth International Conference on Computer Vision*, pp. 439-446

32. **G. Wei and S. Ma.**, 1994: Implicit and explicit camera calibration: Theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16(5), pp. 469-480

33. **G.-Q. Wei and S. De Ma.**, 1994: Implicit and explicit camera calibration : Theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, pp. 469-480

34. **H. Zhang, K.-Y. K. Wong, and G. Zhang.**, 2005: Camera calibration with spheres: Linear approaches. *In Proceedings of the International Conference on Image Processing.* Vol. (2), pp. 1150–1153.

35. **H. Zhang, K.-Y. K. Wong, and G. Zhang.,** 2007: Camera Calibration from Images of Spheres. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29 (3), pp. 499-502.

36. **H. Zollner and R. Sablatnig.,** "Pattern and Image Processing Group (PRIP)"., [Online]. Available: www.prip.tuwien.ac.at/people/sab/papers/oagm04b.pdf

37. **I. Sutherland.**, 1974: Three-dimensional data input by tablet. *In Proceedings of the IEEE*, Vol. 62, pp. 453-461

38. **J. G. Fryer and D. C. Brown.**, 1986: Lens distortion for close-range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, Vol. 52, pp. 51-58

39. **J. Heikkila and O. Silven.**, 1997: A four-step camera calibration procedure with implicit image correction. *In Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, pp. 1106

40. **J. Heikkila.**, 2000: Geometric Camera Calibration Using Circular Control Points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(10), pp. 1066-1077

41. **J. J. Craig.,** 1986: Introduction to Robotics, Addison-Wesley Publishing Company Inc, USA

42. **J. –M Lavest, M. Viala, and M. Dhome.**, 1998: Do we really need an accurate calibration pattern to achieve a reliable camera calibration? *In European Conference on Computer Vision*, Vol. I, pp. 158-174

43. **J. Salvi, X. Armangue, and J. Batlle.**, 2002: A comparative review of camera calibrating methods with accuracy evaluation. *The Journal of the Pattern Recognition Society*, Vol. 35, pp. 1617-1635

44. **J. Weng, P. Cohen, and M. Herniou.**, 1990: Calibration of stereo cameras using a nonlinear distortion model. *In Proceedings of the 10th International Conference on Pattern Recognition*, Vol. 1, pp. 246-253

45. **J. Weng, P. Cohen, and M. Herniou.**, 1992: Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14(10), pp. 965-980

46. **K. W. Wong.**, 1975: Mathematical formulation and digital analysis in close-range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, Vol. 41, pp. 1355-1373

47. **L. L. Wang and W. Tsai.**, 1991: Camera calibration by vanishing lines for 3-D computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, pp. 370-376

48. **L. M. Lye.**, "Similitude, Modelling and Data Analysis"., [Online]. Available: http://www.engr.mun.ca/~llye/engr9516.html

49. **L. M. Lye.**, 2003: Some applications of statistical design of experiment methodology in civil engineering. In Proceedings of the Annual Canadian Society for Civil Engineering Conference, pp. GCD-115-(1-10)

50. **L. M. Lye.**, 2005: Tools and toys for teaching DOE methodology. In Proceedings of the 33rd Annual Canadian Society for Civil Engineering Conference, pp. GC-113--(1-9)

51. **M. Agrawal and L. S. Davis.,** 2003: Camera calibration using spheres: A semi-definite programming approach. *In Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV'03),* Vol. 2, pp. 782-789.

52. **M. Ito.,** 1991: Robot vision modelling- camera modelling and camera calibration. *Advanced Robotics*, Vol. 5, pp. 321-335

53. **M. Penna.**, 1991: Camera calibration: a quick and easy way to determine the scale factor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, pp. 1240-1245

54. **N. Krouglicof.,** "Industrial Machine Vision"., [Online]. Available: http://www.engr.mun.ca/~nick/

55. **O. D. Faugeras and G. Toscani.**, 1986: The calibration problem for stereo. *Proceedings of the IEEE Computer Vision and Pattern Recognition*, pp. 15-20

56. **O. Faugeras, T. Luong, and S. Maybank.**, 1992: Camera self-calibration: theory and experiments. *In European Conference on Computer Vision*, pp. 321-334

57. **O.D. Faugeras and G. Toscani.**, 1986: The calibration problem for stereo. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 15-20

58. **O.D. Faugeras and G. Toscani.**, 1987: Camera Calibration for 3-D Computer Vision. *Proc. Int'l Workshop Industrial Applications of Machine Vision and Machine Intelligence*, pp. 240-247

59. **Q.-T. Luong and O. Faugeras.**, 1997: Self-calibration of a moving camera from point correspondences and fundamental matrices. *The International Journal of Computer Vision*, Vol. 22(3), pp. 261-289

60. **R. H. Myers and D. C. Montgomery.**, 2009: Response Surface Methodology: Process and Product Optimization Using Designed Experiments, 3rd Edition, John Wiley & Sons, Inc.

61. **R. Hartley and A. Zisserman.**, 2000: *Multiple view geometry in computer vision.* Cambridge University Press

62. **R. Hartley and A. Zisserman.**, 2003: *Multiple View Geometry in Computer Vision.* Cambridge University Press. pp. 155-157

63. **R. I. Hartley.**, 1994: An algorithm for self calibration from several views. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 908-912

64. **R. K. Lenz and R. Y. Tsai.**, 1988: Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10 (5), pp. 713-720

65. **R. L. Mason, R. F. Gunst, and J. L. Hess.**, 2003: Statistical Design and Analysis of Experiments with Applications to Engineering and Science, 2nd Edition, John Wiley & Sons, Inc.

66. **R.Y. Tsai.**, 1987: A Versatile Camera Calibration Technique for High-Accuracy 3-D Machine Vision Metrology Using Off-the-Self TV Cameras and Lenses. *IEEE Journal of Robotics and Automation*, vol.3, no. 4, pp. 323-344

67. **S. Chen and W. Tsai.**, 1990: A systematic approach to analytic determination of camera parameters by line features. *Pattern Recognition*, Vol. 23(8), pp. 859-877

68. **S. Ganapathy.**, 1984: Decomposition of transformation matrices for robot vision. *Pattern Recognition Letters*, Vol. 2, pp. 401-412

69. **S. Ganapathy.**, 1984: Decomposition of transformation matrices for robot vision. *In Proceedings of IEEE International Conference on Robotics and Automation*, pp. 130-139

70. **S. J. Maybank and O. D. Faugeras.**, 1992: A theory of self-calibration of a moving camera. *The International Journal of Computer Vision*, Vol. 8(2), pp. 123-152

71. **S. N. R. Wijewickrema, A. P. Paplinski, and C. E. Esson.**, 2006: Reconstruction of Spheres using Occluding Contours from Stereo Images. *In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, pp. 151-154.

72. **S.W. Shih, Y.P. Hung, and W.S. Lin.,** 1993: Accurate linear technique for camera calibration considering lens distortion by solving an eigenvalue problem. *Optical Engineering*, Vol. 32 (1), pp. 138–149

73. **T. A. Clarke and J. G. Fryer.,** 1998: The development of camera calibration methods and models. *The Photogrammetric Record*, Vol. 16(91), pp. 51-66

74. **T. Echigo.,** 1989: A camera calibration technique using three sets of parallel lines. *International Workshop on Industrial Applications of Machine Intelligence and Vision (MIV-89)*, pp. 151-156

75. **T. Melen.,** 1994: Geometrical modelling and calibration of video cameras for underwater navigation. Doctoral dissertation, Norwegian University of Science and Technology, Trondheim, Norway

76. **W. Faig.,** 1975: Calibration of close-range photogrammetry systems: Mathematical formulation. *Photogrammetric Engineering and Remote sensing*, Vol. 41(12), pp. 1479-1486

77. **W. G. Cochran and G. M. Cox.,** 1992: Experimental Designs, 2nd Edition, Wiley, New York.

78. **W. Sun and J. R. Cooperstock.,** 2005: Requirements for camera calibration: Must accuracy come with a high price? *In Proceedings of the Seventh IEEE Workshop on Applications of Computer Vision (WACV/MOTION'05), IEEE Computer Society*, pp. 356-361

79. **X. Ying and H. Zha.**, 2006: A Novel Linear Approach to Camera Calibration from Sphere Images. *In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, Vol. (1), pp. 535-538.

80. **Y. I. Abdel-Aziz and H. M. Karara.**, 1971: Direct linear transformation into object space coordinates in close-range photogrammetry. *In Proceedings of Symposium on Close-Range Photogrammetry*, pp. 1-18

81. **Y. Liu, T.S. Huang, and O.D. Faugeras.**, 1990: Determination of camera location from 2-D to 3-D line and point correspondences, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, pp. 28-37

82. **Z. Hong and J. Yang.**, 1993: An algorithm for camera calibration using a three-dimensional reference point. *Pattern Recognition*, Vol. 26, pp. 1655-1660

83. **Z. Zhang.**, 2000: A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22(11), pp. 1330-1334

84. **Z. Zhang.**, 2002: Camera Calibration with One-Dimensional Objects. *Microsoft Research, Microsoft Corporation, Technical Report*, MSR-TR-2001-120

85. **Z. Zhang.**, 2008: A flexible new technique for camera calibration. *Microsoft Research, Microsoft Corporation, Technical Report*, MSR-TR-98-71

# Appendix A

# Response Surface Methodology

This section briefly describes the use and importance of Response surface methodology in the analysis of designed experiments. Response surface methodology, or RSM, is a collection of mathematical and statistical techniques useful for the modeling and analysis of problems in which a response of interest is influenced by several variables and the objective is to optimize this response. A more detailed explanation of RSM can be obtained in (Montgomery, 2008). The uses of RSM are:

1) To determine the factor levels that will simultaneously satisfy a set of desired specification (e.g. model calibration).
2) To determine the optimum combination of factors that yield a desired response and describe the response near the optimum.
3) To determine how a specific response is affected by changes in the level of the factors over the specified levels of interest.

4) To achieve a quantitative understanding of the system behavior over the region tested.
5) To find conditions for process stability, i.e. insensitive spot (robust condition).
6) To replace a more complex model with a much simpler second order regression model for use within a limited range → replacement models, meta models, or surrogate models.

In most RSM problems, the form of the relationship between the response and the independent variables is unknown. Thus, the first step in RSM is to find a suitable approximation for the true functional relationship between y and the set of independent variables. Usually, a low-order polynomial in some region of the independent variables is employed. If the response is well modeled by a linear function of the independent variables, then the approximating function is the first-order model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad \text{(A.1)}$$

If there is a curvature in the system, then a polynomial of higher degree must be used, such as the second order model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^{\,2} + \sum_{i<j} \sum_{j=1}^{k} \beta_{ij} x_i x_j + \varepsilon \qquad \text{(A.2)}$$

Almost all RSM problems use one or both of these models.

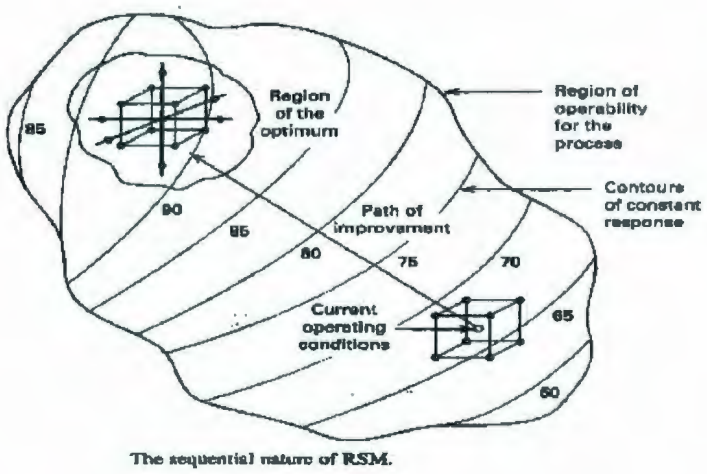



Figure A.1: The sequential nature of RSM

RSM is a sequential procedure. When we are at a point on the response surface that is far from the optimum, i.e. far from the peak such as the current operating conditions in Figure A.1, there is little curvature in the system, then a first-order model will be appropriate. The objective here is to lead the experimenter rapidly and efficiently to the general vicinity of the optimum. Once the region of the optimum has been found, a more elaborate model such a second order model may be employed, and an analysis performed to locate the optimum.

The method of least squares is used to estimate the parameters in the approximating polynomials. The response surface analysis is then performed using the fitted surface. The model parameters can be estimated most effectively if proper experimental designs are used to collect the data. Designs for fitting response surfaces are called response surface designs. Two very useful and popular experimental designs that allow a second order model to be fit are the:

- **Central Composite Design (CCD)**
- **Box-Behnken Design (BBD)**

Both designs are built up from simple factorial or fractional factorial designs.

The Figure A.2 shows a 3-D view of a CCD:

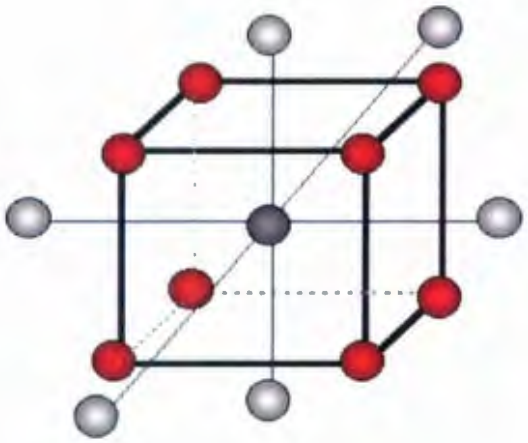


Figure A.2: 3-D view of a CCD

When selecting a response surface design, some of the features of a desirable design are as follows:

1. Provides a reasonable distribution of data points (and hence information) throughout the region of interest.
2. Allows model adequacy, including lack of fit, to be investigated.
3. Allows experiments to be performed in blocks.
4. Allows designs of higher order to be built up sequentially.

5. Provides an internal estimate of error.
6. Provides precise estimates of the model coefficients.
7. Provides a good profile of the prediction variance throughout the experimental region.
8. Provides reasonable robustness against outliers or missing values.
9. Does not require a large number of runs.
10. Does not require too many levels of the independent variables.
11. Ensures simplicity of calculation of the model parameters.

# Appendix B

# CCD Design

This appendix presents a brief overview of the Central Composite Design (CCD). For a detailed description refer to (Montgomery, 2008). A CCD is commonly used to fit a second order model. This is the most popular class of design used for building second order response surface models.

Generally, the CCD consists of a $2^k$ factorial (or fractional factorial of resolution V) with $n_F$ factorial runs ($n_F$ is the number of points used in the factorial portion of the design), 2k axial or star runs, and $n_c$ center runs where k equals the number of factors. Figure B.1 represents central composite design where a $2^k$ design is augmented with four axial runs. The small circles in the figure represent the points at which the experiment is run. For k=2, the design forms a square and for k=3, it forms a cube.

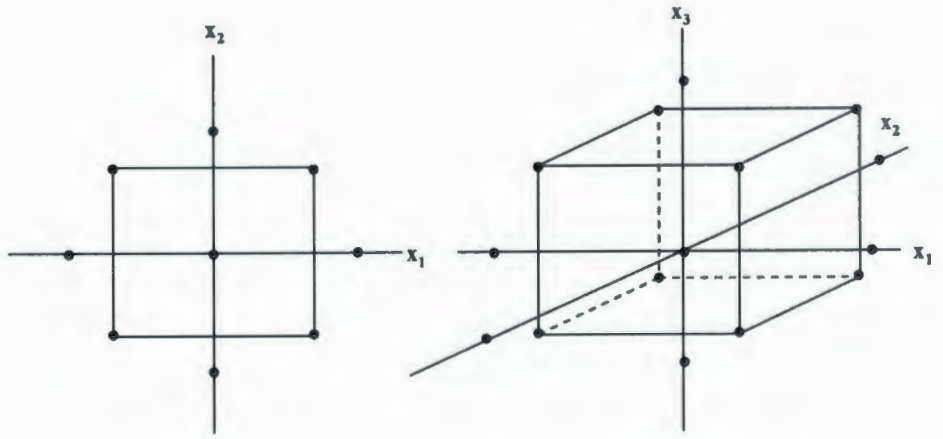



Figure B.1: Central Composite Designs (a) For k=2 (b) For k=3

The practical deployment of a CCD often arises through sequential experimentation. That is, if the curvature is important (exhibits lack of fit), the two level design ($2^k$) can be augmented with axial runs allowing the quadratic terms to be incorporated into the first order model, to obtain a central composite design. That are two parameters in the design that must be specified: the distance $\alpha$ of the axial runs from the design center and the number of center points $n_c$.

The features of CCD are (Lye, [Online]):

- Each factor varies over five levels
- Typically smaller than Box-Behnken designs
- Built upon two-level factorials or fractional factorials of Resolution V or greater

- Can be done in stages → factorial + centerpoints + axial points
- Rotatable

Figure B.2 explains how a CCD is generated:

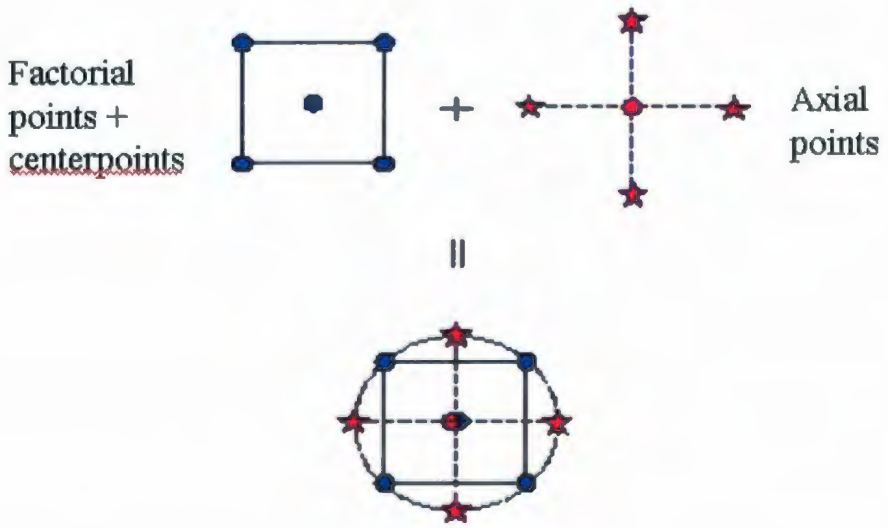



Figure B.2: Illustration of generating a Central Composite Design

Axial points are points on the coordinate axes at distances "a" from the design center; that is, with coordinates. The '$\alpha$' value is usually chosen so that the CCD is rotatable.

**Rotatabilty**: Rotatability is a desirable property for response surface designs, i.e. quadratic model designs (Lye, [Online]). It is important for the second order model to provide good predictions throughout the region of interest (Montgomery, 2008). A design is rotatable if the variance of the predicted response at any point $x$ depends only on the distance of $x$ from the design center point. A design with this property can be rotated around its center point without changing the prediction variance at $x$. That means, for a

rotatable design the variance of the predicted response is constant at all points that are equidistant from the center of the design. Figure B.3 illustrates the three types of central composite designs for two factors. The CCC (Central Composite Circumscribed) explores the largest process space and the CCI (Central Composite Inscribed) explores the smallest process space. Both the CCC and CCI are rotatable designs, but the CCF (Central Composite Face- Centered) is not. In the CCC design, the design points describe a circle circumscribed about the factorial square. For three factors, the CCC design points describe a sphere around the factorial cube.

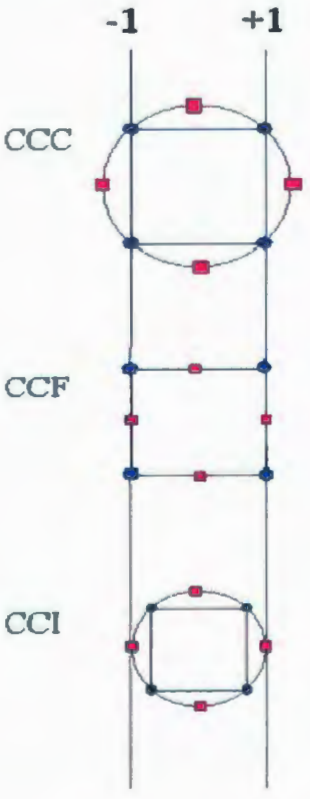



Figure B.3: Three types of Central Composite Designs for k=2

# Appendix C

# Model Adequacy Checking

This appendix describes how a model adequacy checking was performed for the model developed in Chapter 5, through the graphical analysis of the residuals. It was already explained that any violations of the model assumptions (normality of residuals, constant variance, and independence) and model inadequacies can be easily investigated by the examination of residuals, especially by the graphical analysis of the residuals (Montgomery, 2008). In general, if the model is adequate, the residuals should contain no obvious patterns.

A check of normality assumption could be made by constructing a normal probability plot of the residuals. If the error distribution is normal then the plot should resemble a straight line. A check of constant variance assumption could be made by constructing a plot of the residuals versus fitted or predicted values. If the model is correct and the assumptions are satisfied, the residuals should be structureless; i.e. the plot should not reveal any obvious pattern. In particular, the residuals should be unrelated

to any other variable including the predicted response. Plotting of residuals in run/time sequence will check the independence assumption. This plot is helpful in detecting the correlation between the residuals. Proper randomization of the experiment is an important step in obtaining independence. The presence of outliers can also be checked for proving the adequacy of the model. A residual that is very much larger than any of the others in the normal probability plot or in the plot of outlier T or externally studentized residuals is often called an outlier. The presence of one or more outliers can seriously affect the analysis of variance. Figures C.1a to C.7 shows the results of analysis of residuals for model assumptions.

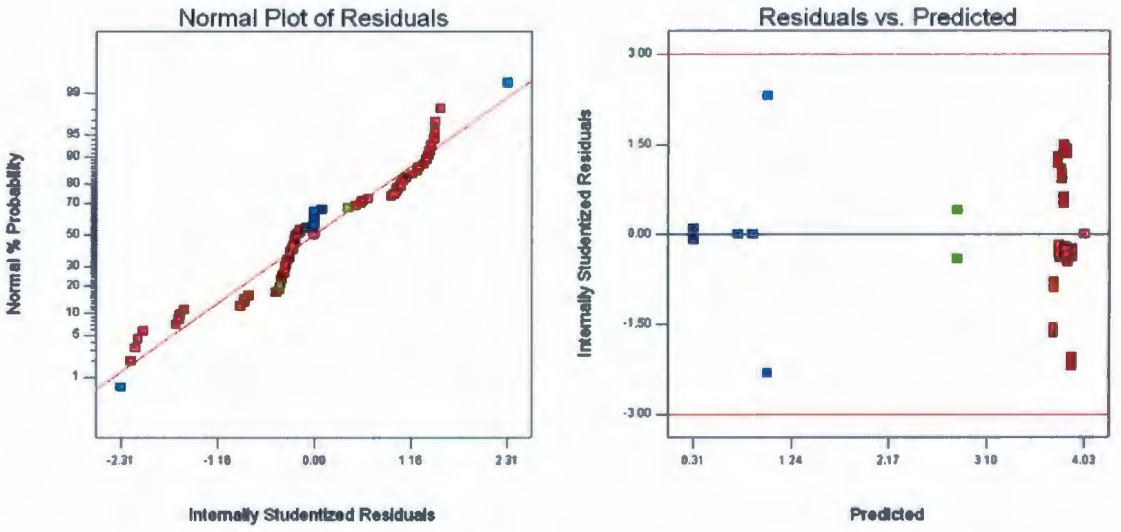



Figure C.1: (a) Normal probability plot of residuals (b) Plot of residuals versus predicted

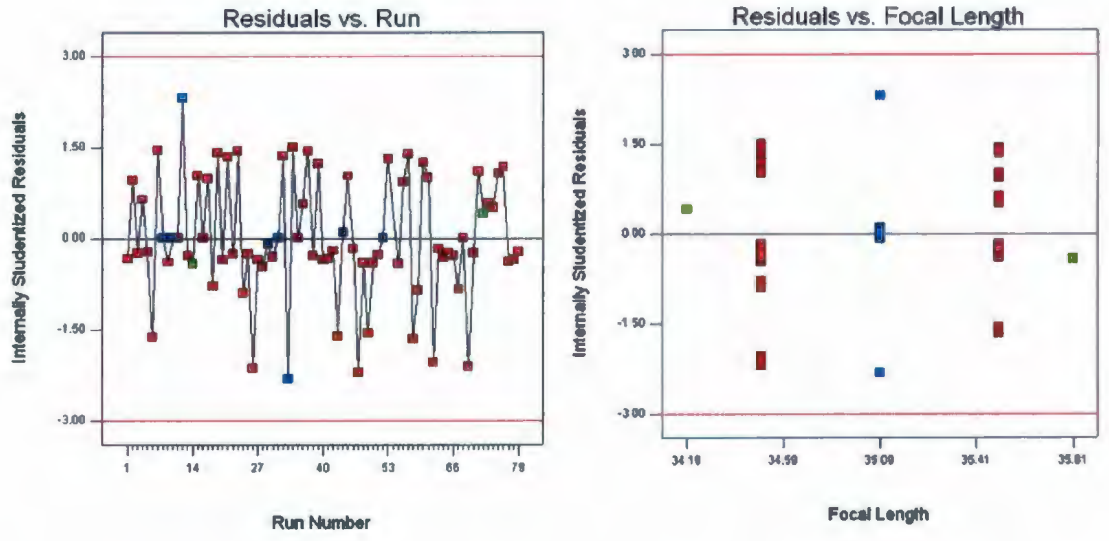



Figure C.2: (a) Plot of residuals versus run (b) Plot of residuals versus factor A

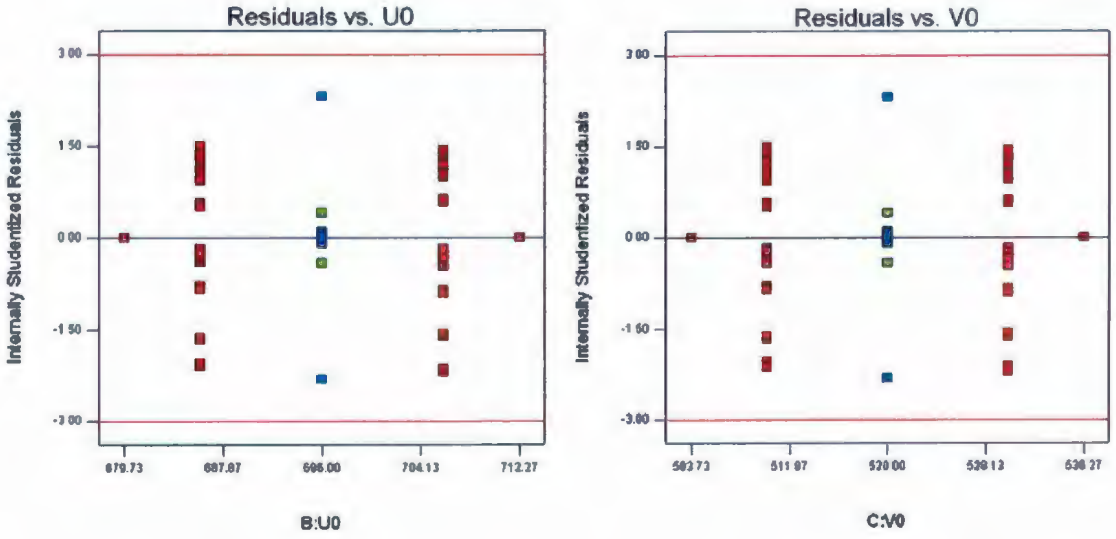



Figure C.3: (a) Plot of residuals versus factor B (b) Plot of residuals versus factor C

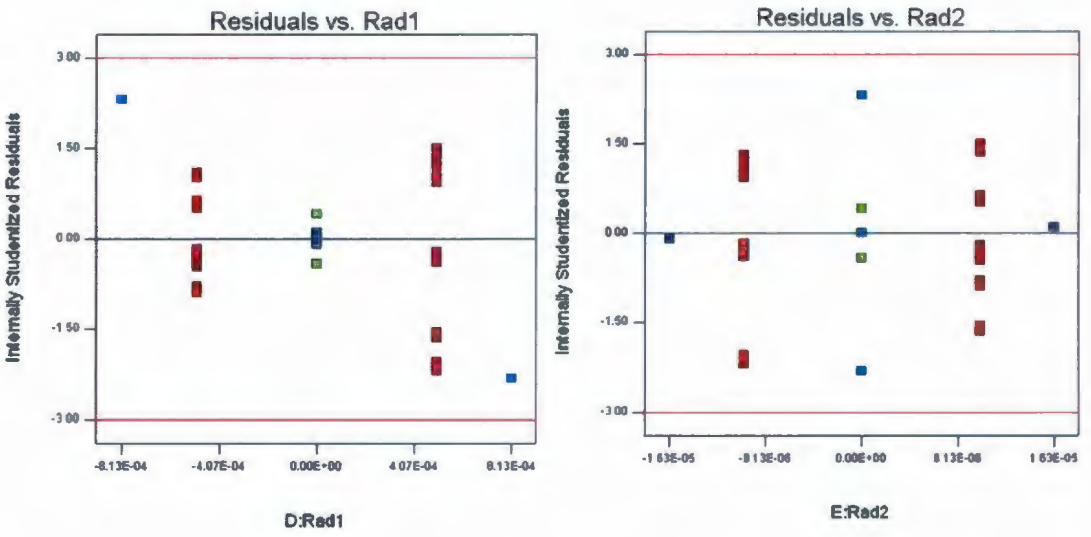



Figure C.4: (a) Plot of residuals versus factor D (b) Plot of residuals versus factor E

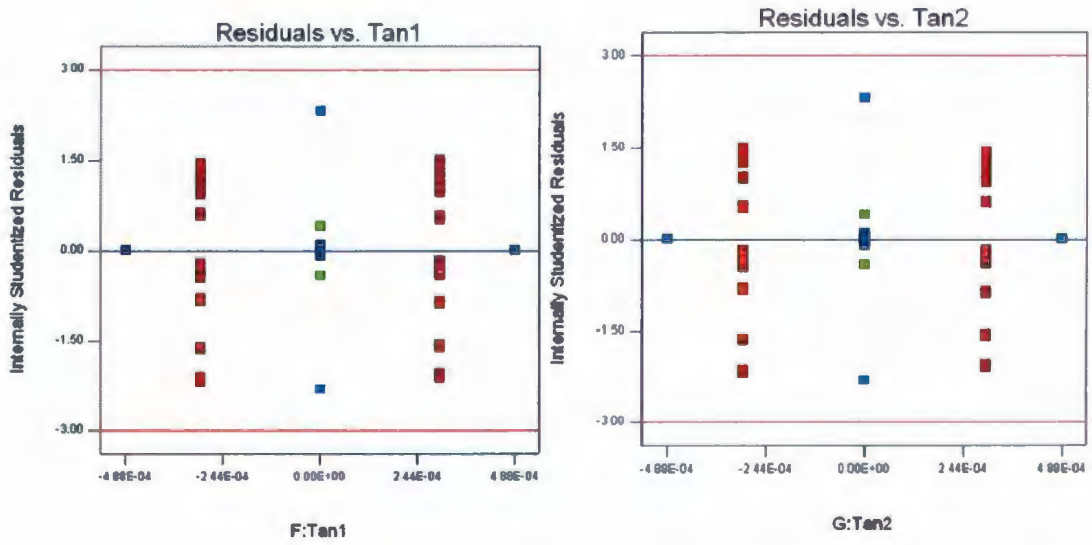



Figure C.5: (a) Plot of residuals versus factor F (b) Plot of residuals versus factor G

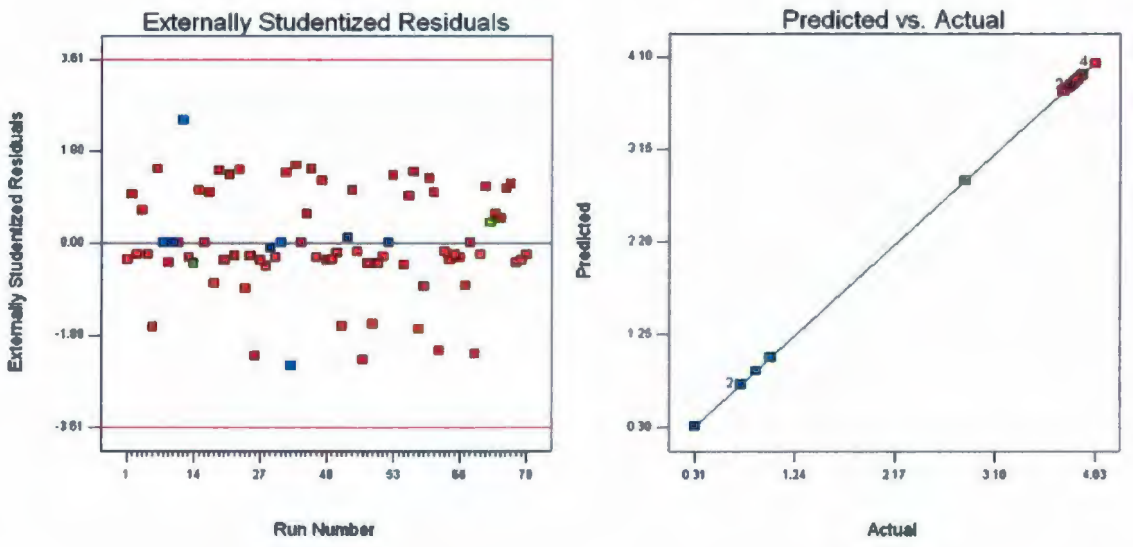



Figure C.6: (a) Plot of externally studentized residuals versus run number (b) Plot of predicted versus actual

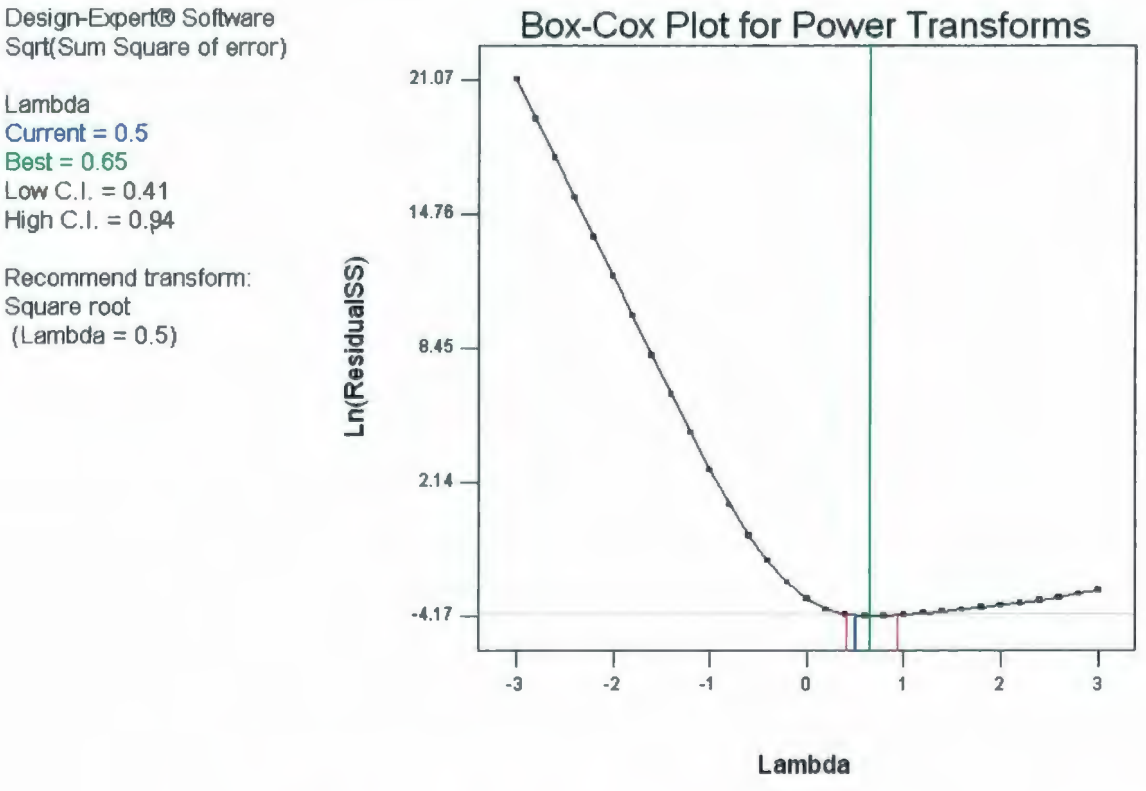



Figure C.7: Box-Cox Plot for power transforms

Figure C.1a shows almost a straight line fit. This indicates that the residuals are approximately normally distributed. In general, moderate departures from normality are of little concern in the fixed effects analysis of variance. Figure C.1b plots the residuals against the predicted values showing a nice scatter without any systematic pattern. This indicates a constant variance. The points show random scattering or pattern in the Figure C.2a where the residuals are plotted against run order or time. Also Figures C.2b to C.5b do not show any obvious pattern, indicating that there is no particular contribution of an independent factor that is not accounted for in the model. This implies the independence assumption. Figure C.6a shows any significant outliers. The outliers lie between the limits and no visible pattern can be seen. The plot of predicted versus actual in Figure C.6b shows consistency. The Box-Cox plot in Figure C.7 shows a square root transformation of the model as recommended by the plot before. All plots from Figure C.1a to C.7 explained before had been obtained by undertaking the square root transformation on the model recommended by the Box-Cox plot as the transformation gave good results compared to no transformation of the model. Therefore from the above graphs and the underlying statements the model adequacy was proven.

# Appendix D

# Newton's Method

This appendix explains the numerical technique namely, Newton's method for solving the nonlinear systems of equations developed from the geometrical model considered in Chapter 6 (Krouglicof, [Online]).

Generally, if there are n unknown parameters to be determined then a minimum of n independent equations are required for a solution. But if the equations are nonlinear in nature, then an analytical approach is not always feasible and a numerical technique must be applied to solve the equations. Newton's method for solving nonlinear system of equations is a simple numerical method for finding the zeros or roots of a nonlinear differential function. The method works in an iterative manner. The method can be described as follows:

The system of n nonlinear equations is represented by the vector $F(\bar{\theta})$:

$$F(\bar{\theta}) = \begin{bmatrix} f_1(\bar{\theta}) \\ f_2(\bar{\theta}) \\ \cdots \\ f_n(\bar{\theta}) \end{bmatrix} \tag{D.1}$$

where the vector $\bar{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdots \\ \theta_n \end{bmatrix}$ which represents the n unknown parameters.

The first order Taylor series approximation of $F(\bar{\theta})$ about an initial parameter vector $\bar{\theta}_0$ is given by:

$$F(\bar{\theta}) = F(\bar{\theta}_0) + J(\bar{\theta}_0) \cdot (\bar{\theta} - \bar{\theta}_0) \tag{D.2}$$

where $J(\bar{\theta}_0)$ is the Jacobian of $F(\bar{\theta})$ evaluated at $\bar{\theta}_0$ given by:

$$J(\bar{\theta}) = \begin{bmatrix} \frac{\delta f_1(\bar{\theta})}{\delta \theta_1} & \frac{\delta f_1(\bar{\theta})}{\delta \theta_2} & \cdots & \frac{\delta f_1(\bar{\theta})}{\delta \theta_n} \\ \frac{\delta f_2(\bar{\theta})}{\delta \theta_1} & \frac{\delta f_2(\bar{\theta})}{\delta \theta_2} & \cdots & \frac{\delta f_2(\bar{\theta})}{\delta \theta_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\delta f_n(\bar{\theta})}{\delta \theta_1} & \frac{\delta f_n(\bar{\theta})}{\delta \theta_2} & \cdots & \frac{\delta f_n(\bar{\theta})}{\delta \theta_n} \end{bmatrix} \tag{D.3}$$

The Jacobian is defined as the matrix of partial derivatives of each of the n equations contained in $F(\bar{\theta})$ evaluated with respect to each of the unknown parameters contained in $\bar{\theta}$. Since the objective of Newton's method is to find the zeros or roots of $F(\bar{\theta})$, Equation (D.2) can be written as:

$$F(\bar{\theta}) = 0 = F(\bar{\theta}_0) + J(\bar{\theta}_0) \cdot (\bar{\theta} - \bar{\theta}_0) \tag{D.4}$$

Solving for $\bar{\theta}$ from Equation (D.4) yields:

$$0 = J(\bar{\theta}_0)^{-1} \cdot F(\bar{\theta}_0) + (\bar{\theta} - \bar{\theta}_0) \tag{D.5}$$

For simplicity, $(\bar{\theta} - \bar{\theta}_0)$ can be represented by a vector $\bar{\beta}$, i.e. $\bar{\beta} = (\bar{\theta} - \bar{\theta}_0)$. Thus by substituting $\bar{\beta}$, Equation (D.5) can be written as

$$0 = J(\bar{\vartheta}_0)^{-1} . F(\bar{\theta}_0) + \bar{\beta} \tag{D.6}$$

The vector $\bar{\beta}$ can be evaluated by solving the following linear system of equations:

$$J(\bar{\theta}_0) \cdot \bar{\beta} = -F(\bar{\theta}_0) \tag{D.7}$$

A new estimate of $\bar{\theta}$ is computed by adding $\bar{\beta}$ and $\bar{\theta}_0$ as follows:

$$\bar{\theta} = \bar{\theta}_0 + \bar{\beta} \tag{D.8}$$

Being an iterative technique, the algorithm continues until the specified accuracy criterion is attained or until the maximum number of iterations is exceeded (i.e. the algorithm fails). A suitable accuracy criterion can be defined as follows:

$$\left\| F(\bar{\theta}) \right\| = \sqrt{f_1(\bar{\theta})^2 + f_2(\bar{\theta})^2 + \ldots + f_n(\bar{\theta})^2} \tag{D.9}$$

Basically the algorithm will successfully converge to a solution when the Euclidean norm is less than some threshold, $\xi$.

A pseudo-code implementation of the Newton's method can be written as follows:

$\overline{\theta}_0$ = initial estimate of root

i = 1; ($i$ = number of iterations)

$\xi$ = required accuracy

$nrm$ = 1; ($nrm$ = Euclidean norm)

$n\_max$ = maximum number of iterations allowed

while ($nrm > \xi$) and ($i<n\_max$)

{

$$nrm = sqrt\left(f_1\left(\overline{\theta}\right)^2 + f_2\left(\overline{\theta}\right)^2 + \ldots + f_n\left(\overline{\theta}\right)^2\right)$$

$$\overline{\beta} = invJ\left(\overline{\theta}_0\right)\left(-F\left(\overline{\theta}_0\right)\right)$$

$$\overline{\theta} = \overline{\theta}_0 + \overline{\beta}$$

$$i = i + 1$$

}

# Appendix E

# Nonlinear Least-Squares Analysis

This appendix explains the technique of Nonlinear Least-Squares Analysis used in Chapter 6 for the problem of fitting a nonlinear model (Krouglicof, [Online]).

Nonlinear least-square analysis refers to the problem of fitting a nonlinear model to a set of experimental data points. Generally, if there are n unknown parameters to be determined then a minimum of n independent equations are required for a solution. But in this case there are m unknown parameters and n equations; i.e. one per data point. The method also works in an iterative fashion. The method can be described as follows:

The system of n nonlinear equations is represented by the vector $F(\bar{\theta})$:

$$F(\bar{\theta}) = \begin{bmatrix} f_1(\bar{\theta}) \\ f_2(\bar{\theta}) \\ \cdots \\ f_n(\bar{\theta}) \end{bmatrix} \tag{E.1}$$

where the vector $\bar{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdots \\ \theta_m \end{bmatrix}$ which represents the m unknown parameters.

The first order Taylor series approximation of $F(\bar{\theta})$ about an initial parameter vector $\bar{\theta}_0$ is given by:

$$F(\bar{\theta}) = F(\bar{\theta}_0) + J(\bar{\theta}_0) \cdot (\bar{\theta} - \bar{\theta}_0) \tag{E.2}$$

where $J(\bar{\theta}_0)$ is the Jacobian of $F(\bar{\theta})$ evaluated at $\bar{\theta}_0$.

The Jacobian is defined as the matrix of partial derivatives of each of the n equations contained in $F(\bar{\theta})$ evaluated with respect to each of the m unknown parameters contained in $\bar{\theta}$. As in the case of Newton's method, the objective of nonlinear least squares analysis is also to find the zeros or roots of $F(\bar{\theta})$, and Equation (E.2) can be written as:

$$F(\bar{\theta}) = 0 = F(\bar{\theta}_0) + J(\bar{\theta}_0) \cdot (\bar{\theta} - \bar{\theta}_0) \tag{E.3}$$

Equation (D.3) can be simplified by defining a vector, $\bar{\beta}$, given by:

$$\bar{\beta} = (\bar{\theta} - \bar{\theta}_0)$$

Substituting $\bar{\beta}$ in Equation (E.3) yields:

$$0 = F(\bar{\theta}_0) + J(\bar{\vartheta}_0)\bar{\beta} \tag{E.4}$$

If the experimental data is subject to measurement error and the system of equations is overdetermined (i.e. n>m), then Equation (E.4) cannot be satisfied explicitly for all points. An error term, $\bar{\varepsilon}$, must be introduced to account for this discrepancy:

$$(\bar{\varepsilon}) = F(\bar{\theta}_0) + J(\bar{\vartheta}_0)\bar{\beta} \tag{E.5}$$

The sum of the squares of the individual error is an appropriate error criterion, q, which can be readily minimized. The sum of the squares of errors can be represented in matrix form as follows:

$$q = (\bar{\varepsilon})^T.(\bar{\varepsilon}) = \left(F(\bar{\theta}_0) + J(\bar{\theta}_0)(\bar{\beta})\right)^T.\left(F(\bar{\theta}_0) + J(\bar{\theta}_0)(\bar{\beta})\right) \tag{E.6}$$

In order to minimize q, the partial derivative of Equation (E.5) with respect to each of the unknown parameters must be calculated and set equal to zero. After some manipulation this yields the following set of normal equations:

$$(\bar{\beta}) = -\left(J(\bar{\theta}_0)^T.J(\bar{\theta}_0)\right)^{-1}.J(\bar{\theta}_0)^T.F(\bar{\theta}_0) \tag{E.7}$$

Once again the algorithm proceeds in an iterative manner. A new estimate of $\bar{\theta}$ is computed by adding $\bar{\beta}$ and $\bar{\theta}_0$ as follows:

$$\bar{\theta} = \bar{\theta}_0 + \bar{\beta} \tag{E.8}$$

A pseudo-code implementation of the Nonlinear Least-Squares Analysis can be written as follows:

n = Number of data points

m = Number of unknown parameters

$\bar{\theta}_0$ = initial estimate of root

$\xi$ = required accuracy

cnt =0; (cnt = iteration counter)

*nrm* = 1; (*nrm* = Euclidean norm)

*n_max* = maximum number of iterations allowed

while ($nrm > \xi$) and *cnt*<*n_max*)

{

$\bar{\beta} = -1 * inv(J.'*J) * J.'*F$

nrm = 0

for i = 1:m

nrm = nrm + $\beta(i)^{\wedge 2}$

end

$\bar{\theta} = \bar{\theta}_0 + \bar{\beta}$

*cnt* = cnt+ 1

}

# List of Publications Based on this Thesis

- Swapna, P., Krouglicof, N., and Gosine, R. (2009). The Question of Accuracy with Geometric Camera Calibration. *The 22nd Canadian Conference on Electrical and Computer Engineering,* May3-6, St. John's, NL, Canada.

- Swapna, P., Krouglicof, N., and Gosine, R. (2008). The Problem with Geometric Camera Calibration: Analysis Using Design of Experiments. *The Eighteenth Annual Newfoundland Electrical and Computer Engineering Conference*, Nov 6, St. John's, NL, Canada.

- Swapna, P., Krouglicof, N., and Gosine, R. (2007). Estimation of Accuracy in Geometric Camera Calibration. *The Seventeenth Annual Newfoundland Electrical and Computer Engineering Conference*, Nov 8, St. John's, NL, Canada.