**IB Math Analysis and Approaches (HL)**

**Extended Essay**

**May 2024 Session**

---

# Mathematical Techniques and Applications of Camera Calibration

---

**Research Question:** What mathematical techniques can be employed to develop highly accurate camera models, and what are their real-world applications where these models prove valuable?

**Word Count:** 222 words

# Contents

# 1 Introduction

Camera calibration is an important process in computer vision and computer graphics which involves determining the parameters of a camera. The knowledge of these parameters are essential, because it allows us to create a mathematical model which accurately describes the camera. Without a well-calibrated camera, images captured may suffer from inaccuracies and distortion, making calibration an indispensable step in a wide array of applications.

## 1.1 Problem Statement

While manufacturers of cameras often report parameters of cameras, such as the nominal focal length and pixel sizes of their camera sensor, these figures are typically approximations which can vary from camera to camera, particularly in consumer-grade cameras. As such, the use of these estimates by manufacturers are unsuitable in developing camera models for applications requiring high accuracy. Combined with the potential for manufacturing defects as well as unknown lens distortion coefficients further necessitates the need for a reliable method for determining the parameters of a camera.

Camera calibration emerges as the answer to these problems, allowing us to create very accurate estimates for the parameters of a camera. As such, it is important that we ac

Photogrammetry, as a comprehensive science, concerns itself with obtaining precise measurements of 3-dimensional physical objects from photographic images.

It was first employed by Prussian architect Albrecht Meydenbauer in the 1860s, who used photogrammetric techniques to create some of the most detailed topographic plans and elevations drawings[1]. Camera calibration borrows techniques from photogrammetry

Today, photogrammetric techniques are used in a multitude of applications spanning diverse fields, including computer vision, topographical mapping, medical imaging, and forensic analysis.

Importance of reserach question

In order to accurately determine the position of 3D points based on data from multiple 2D images, we must have knowledge of the parameters of the camera.

This process of calculating

[1] Albertz, "A Look Back; 140 Years of Photogrammetry," 504–506.

# 2 Overview

The techniques used

There are countless different approaches one could take to calibrate a camera, however they all build upon techniques first described in multiple highly influential papers, most notably Tsai's "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses" and Zhang's "A Flexible New Technique for Camera Calibration".

Calibration techniques can be roughly separated into 3 categories, based on the dimension of the calibration object used[2]:

# 3 Camera Model

A camera model is a projection model that approximates the function of a camera by describing a mathematical relationship between points in 3D space and its projection onto the sensor grid of the camera. In order to accurately model a camera, we must first understand the general workings of a camera.

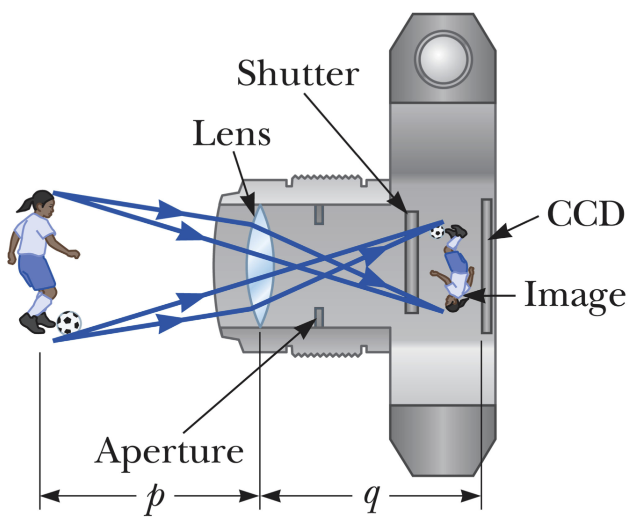


**Figure 3.1:** Lens camera. Adapted from Colton, "Warm-Up Exercise 30."

Modern lens cameras are very complex, as they often contain a series of different shaped lenses. This is an effort to compensate for other undesired effects, such as common phenomenon where the image quality in the center of the camera is better than the edges of image, due to the curvature of the lens. Additionally, they contain various intricate mechanisms, such as

[2] Zhang, "Camera Calibration."

the ability to zoom the camera and other features to alter the image output. However, we can simplify the model of the lens camera by collapsing the mechanisms of the camera into 3 main functional components that are important to the image projection: the lens, the aperture, and the sensor grid (CCD). This simplified model is visualized in Figure 3.1. The lens focuses incoming light rays towards the aperture, before they project inverted onto the sensor grid. However, even this simplified model of a lens camera is too complex to model, as there is no simple mathematical equation which accurately describes the behavior of a lens. As such, we can further simplify our camera model by building upon the pinhole camera model, which is one of the simplest and most commonly used camera models in camera calibration.

A pinhole camera is a simple camera without a lens. Instead, it relies on the use of a tiny hole as the aperture of the camera, and light rays pass through the hole, projecting an inverted image onto the

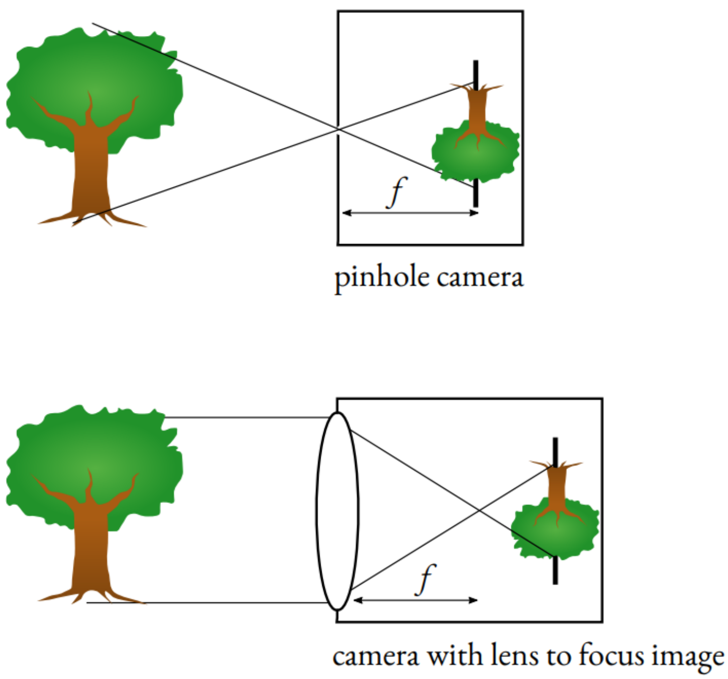


**Figure 3.2:** Difference between a pinhole camera and a lens camera. Adapted from Lê, "Camera Model: Intrinsic Parameters."

There are a few assumptions which are made by the pinhole camera model:

Extremely simple model for imaging geometry Doesn't strictly apply Mathematically convenient acceptable approximation.

- T

The pinhole camera model does not accurately describe the true workings of a camera, as some of the effects that the model fails to account for can be compensated the errors which results from these assumptions are sufficiently small to be neglected if a high quality camera is used. Additionally,

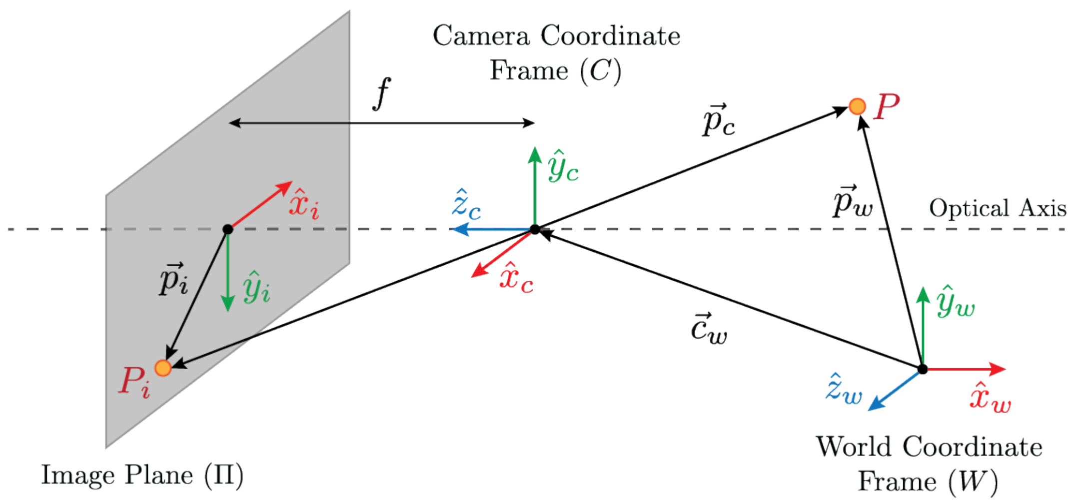


**Figure 3.3:** Pinhole camera model.

## 3.1 Straetgy

## 3.2 Nomenclature

For our camera model, we will establish 3 frames of

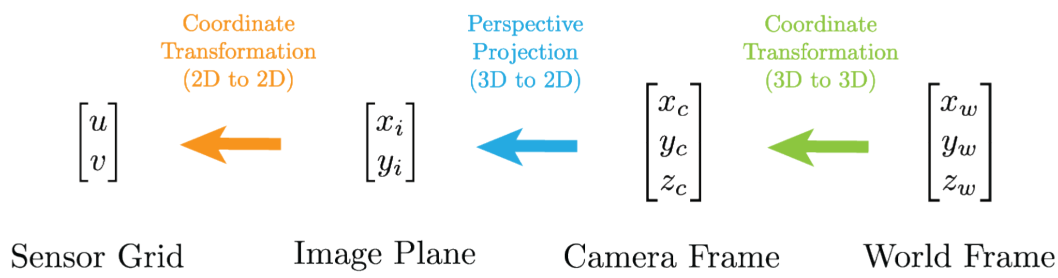


**Figure 3.4:** Coordinate remappings.

# 4 Prerequisites

## 4.1 Homogenous coordinates

## 4.2 Least Squares Solution

# 5 Intrinsic Parameters

Now, we would like to find the intrinsic matrix, $M_{int}$, which relates the positional vector $\vec{p}_c$ of point $P$ in the camera coordinate frame, to its positional vector $\vec{p}_i$ on the image plane. Using $\widetilde{p}_c$ and $\widetilde{p}_i$ to represent the homogenous coordinates of the vectors $\vec{p}_c$ and $\vec{p}_i$ respectively, we can express this mathematically as follows:

$$\widetilde{p}_i = M_{ext}\,\widetilde{p}_c \tag{5.1}$$

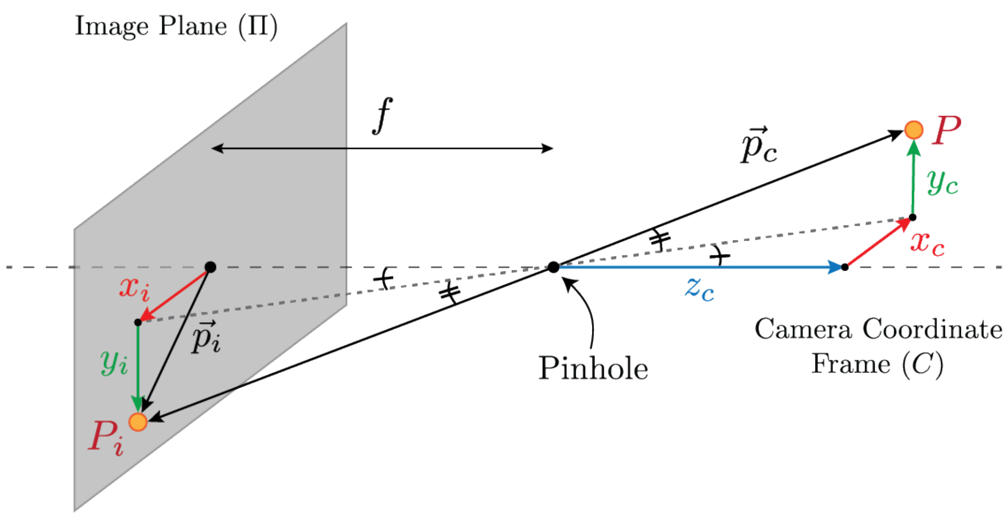


**Figure 5.1:** Perspective projection of the point onto the image plane Π.

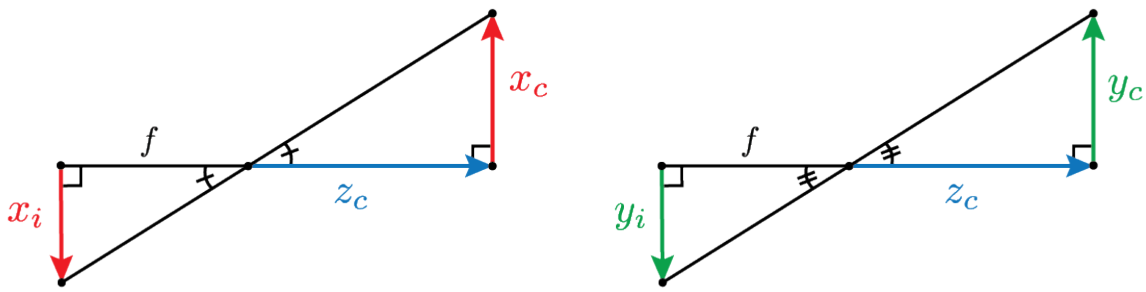


**Figure 5.2:** Similar triangles formed by perspective projection, which relate $x_i$ to $x_c$ and $y_i$ to $y_c$

$$\frac{x_i}{f} = \frac{x_c}{z_c} \implies x_i = f\frac{x_c}{z_c} \tag{5.2a}$$

$$\frac{y_i}{f} = \frac{y_c}{z_c} \implies y_i = f\frac{y_c}{z_c} \tag{5.2b}$$

We can then relate the coordinates of the projection, $(x_i, y_i)$, which are in real-world units, to its position $(u, v)$ in pixels.

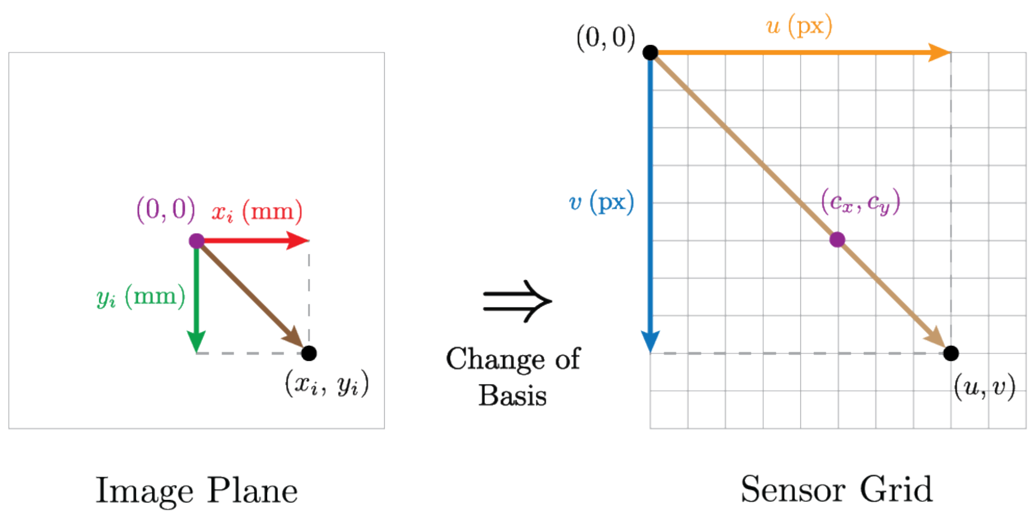


**Figure 5.3:** Conversion from image plane coordinates to sensor grid coordinates

Let $m_x$ and $m_y$ represent the pixel density of the image sensor in the $x$ and $y$ axes of the image sensor plane respectively.

$$u = m_x x_i + c_x \tag{5.3a}$$

$$v = m_y y_i + c_y \tag{5.3b}$$

Replacing $x_i$ and $y_i$ for the result we obtained from 5.2a and 5.2b, we get:

$$u = m_x f \frac{x_c}{z_c} + c_x$$

$$v = m_y f \frac{y_c}{z_c} + c_y$$

Since $m_x$, $m_y$, and $f$ are all unknowns, we can combine the products $m_x f$ and $m_y f$ to $f_x$ and $f_y$ respectively. Under this new scheme, we define $f_x$ and $f_y$ as the horizontal and vertical focal lengths of camera.

$$u = f_x \frac{x_c}{z_c} + c_x \tag{5.4a}$$

$$v = f_y \frac{y_c}{z_c} + c_y \tag{5.4b}$$

We can linearize

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \widetilde{u} \\ \widetilde{v} \\ \widetilde{w} \end{bmatrix} \equiv \begin{bmatrix} z_c u \\ z_c v \\ z_c \end{bmatrix} = \begin{bmatrix} f_x x_c + z_c c_x \\ f_y y_c + z_c c_y \\ z_c \end{bmatrix} = \underbrace{\begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{= M_{int}} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \tag{5.5}$$

# 6 Extrinsic Parameters

Next, we would like to find the extrinsic matrix, $M_{ext}$, which relates the positional vector $\vec{p}_w$ of point $P$ in the world coordinate frame, to its positional vector $\vec{p}_c$ in the camera coordinate frame. Similar to what we did in section 5, we can express this in homogenous coordinates as follows:

$$\widetilde{p}_c = M_{ext}\, \widetilde{p}_w \tag{6.1}$$

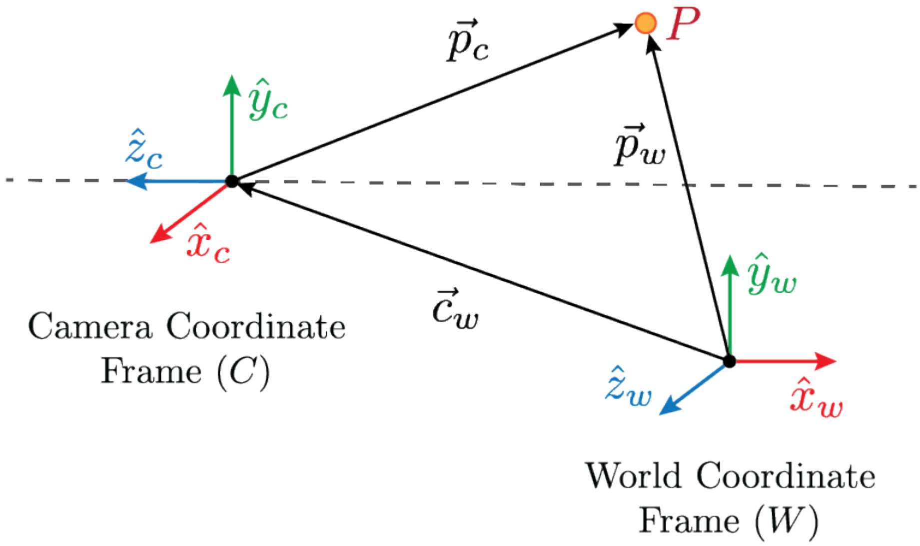


**Figure 6.1:** Coordinate transformation from the world coordinate frame to the camera frame.

For the extrinsic parameters of the camera, we have the position $\vec{c}_w$ of the camera in world coordinates and orientation $R$ of the camera. The orientation, $R$, is a 3x3 rotational matrix:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{6.2}$$

where:

- Row 1: Direction of $\hat{x}_c$ in world coordinate frame.
- Row 2: Direction of $\hat{y}_c$ in world coordinate frame.
- Row 3: Direction of $\hat{z}_c$ in world coordinate frame.

$$\vec{p}_c = R(\vec{p}_w - \vec{c}_w) \tag{6.3a}$$

$$= R\vec{p}_w - R\vec{c}_w \tag{6.3b}$$

$$\vec{p}_c = R\vec{p}_w + \vec{t} \tag{6.4}$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{6.5}$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{= M_{ext}} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{6.6}$$

# 7 Determining the Projection Matrix

Combining the equations $p_c = M_{ext}\, p_w$ and $p_i = M_{int}\, p_c$, we obtain the equation 7.1:

$$\widetilde{p}_i = M_{int}\, M_{ext}\, \widetilde{p}_w = P\, \widetilde{p}_w \tag{7.1}$$

$$\begin{bmatrix} u_n \\ v_n \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \widetilde{w}_n u_n \\ \widetilde{w}_n v_n \\ \widetilde{w}_n \end{bmatrix} \equiv \begin{bmatrix} \widetilde{u}_n \\ \widetilde{v}_n \\ \widetilde{w}_n \end{bmatrix} = \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{= P} \begin{bmatrix} x_w^{(n)} \\ y_w^{(n)} \\ z_w^{(n)} \\ 1 \end{bmatrix} \tag{7.2}$$

$$\widetilde{u}_n = p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14}$$
$$\widetilde{v}_n = p_{21}x_w^{(n)} + p_{22}y_w^{(n)} + p_{23}z_w^{(n)} + p_{24}$$
$$\widetilde{w}_n = p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}$$

$$u_n = \frac{\widetilde{u}_n}{\widetilde{w}_n} = \frac{p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14}}{p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}}$$

$$v_n = \frac{\widetilde{v}_n}{\widetilde{w}_n} = \frac{p_{21}x_w^{(n)} + p_{22}y_w^{(n)} + p_{23}z_w^{(n)} + p_{24}}{p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}}$$

$$u_n(p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}) = p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14}$$

$$v_n(p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}) = p_{21}x_w^{(n)} + p_{22}y_w^{(n)} + p_{23}z_w^{(n)} + p_{24}$$

$$0 = p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14} - p_{31}u_n x_w^{(n)} - p_{32}u_n y_w^{(n)} - p_{33}u_n z_w^{(n)} - p_{34}u_n \tag{7.3a}$$

$$0 = p_{21}x_w^{(n)} + p_{22}y_w^{(n)} + p_{23}z_w^{(n)} + p_{24} - p_{31}v_n x_w^{(n)} - p_{32}v_n y_w^{(n)} - p_{33}v_n z_w^{(n)} - p_{34}v_n \tag{7.3b}$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} x_w^{(1)} & y_w^{(1)} & z_w^{(1)} & 1 & 0 & 0 & 0 & 0 & -u_n x_w^{(1)} & -u_n y_w^{(1)} & -u_n z_w^{(1)} & -u_1 \\ 0 & 0 & 0 & 0 & x_w^{(1)} & y_w^{(1)} & z_w^{(1)} & 1 & -v_n x_w^{(1)} & -v_n y_w^{(1)} & -v_n z_w^{(1)} & -v_1 \\ & & \vdots & & & & & & & \vdots & & \\ x_w^{(n)} & y_w^{(n)} & z_w^{(n)} & 1 & 0 & 0 & 0 & 0 & -u_n x_w^{(n)} & -u_n y_w^{(n)} & -u_n z_w^{(n)} & -u_n \\ 0 & 0 & 0 & 0 & x_w^{(n)} & y_w^{(n)} & z_w^{(n)} & 1 & -v_n x_w^{(n)} & -v_n y_w^{(n)} & -v_n z_w^{(n)} & -v_n \end{bmatrix}}_{=A} \underbrace{\begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix}}_{=p} \tag{7.4}$$

## 7.1 Constrained Least Squares Solution

We have now established a way to solve for the

Now, we need to solve for $Ap = 0$

$$\min_{p} \ \|Ap\|^2 \quad \text{subject to} \quad \|p\|^2 = 1 \tag{7.5}$$

For a given arbitrary vector $v \in \mathbb{R}^n$, the magnitude of the vector, $\|v\|$, is equal to $\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$. As such, we can rewrite the square of the magnitude of vector, $\|v\|^2$, as:

$$\|v\|^2 = v_1^2 + v_2^2 + \cdots v_n^2 = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v^T v$$

Thus, we can replace $\|Ap\|^2$ and $\|p\|^2$ in 7.5 for $p^T A^T Ap$ and $p^T p$ respectively to obtain 7.6.

$$\min_{p} \ \left(p^T A^T Ap\right) \quad \text{subject to} \quad p^T p = 1 \tag{7.6}$$

We can define the loss function:

$$L(p, \lambda) = p^T A^T Ap - \lambda \left(p^T p - 1\right) \tag{7.7}$$

To find the values of vector $p$ such that the $L(p,\lambda)$, we find when derivative of the function with respect to $p$, and find when the derivative is 0.

$$\begin{aligned} \frac{d}{dp} L(p,\lambda) &= \frac{d}{dp} \left[p^T A^T Ap - \lambda \left(p^T p - 1\right)\right] \\ &= 2A^T Ap - 2\lambda p \equiv 0 \end{aligned} \tag{7.8}$$

$$\Rightarrow A^T Ap = \lambda p \tag{7.9}$$

Equation 7.9 in fact takes the form of the well-known eigenvalue problem. which states that for a given matrix $M \in \mathbb{R}^{n \times n}$, determine the eigenvector $x \in \mathbb{R}^n, x \neq 0$ and the eigenvalue $\lambda \in \mathbb{C}$ such that:

$$Mx = \lambda x$$

As such, we can reframe our problem as finding the eigenvector $p$ with the smallest $\lambda$ of the matrix $A^T A$ which minimizes the loss function $L(p,\lambda)$

# 8 Handling Distortion

When constructing our camera model in section 3, we made the assumption that the camera we are using can be approximated using a pinhole camera.

Modern cameras rarely have a single lens. Rather, they typically have compound lens

## 8.1 Symmetrical Radial Distortion

Symmetrical lens distortion refers to

It

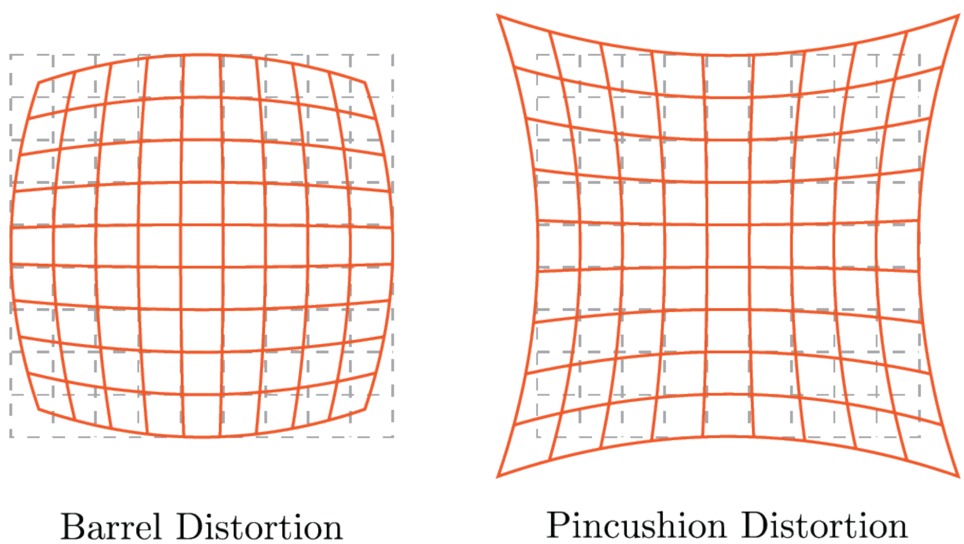


**Figure 8.1:** Types of radial distortion.

## 8.2 Asymmetrical Radial Distortion

## 8.3 Tangential (Decentering) Distortion

of axis

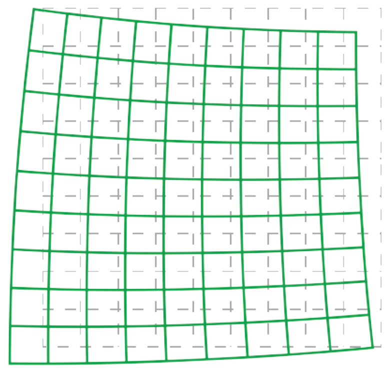


**Figure 8.2:** Tangential (Decentering) Distortion

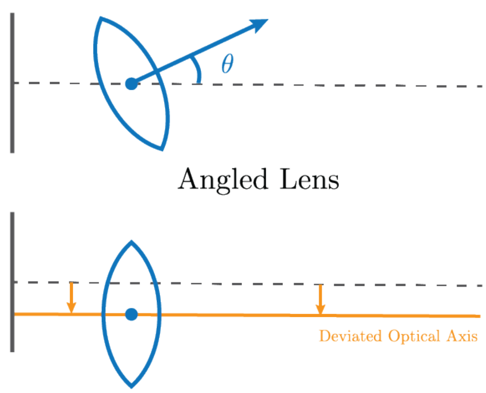


**Figure 8.3:** Reasons why tangential distortion can occur

## 8.4 Brown-Conrady Model

# 9 Applications

# Acknowledgements

I am very grateful to my supervisor Mr. Hoteit for his continual guidance and invaluable pieces of advice during the process of writing this extended essay. I would also like to thank Aditya, for aiding me with the creation of some diagrams used in this essay.

# Bibliography

Albertz, Joerg. "A Look Back; 140 Years of Photogrammetry." *Journal of the American Society for Photogrammetry and Remote Sensing* 73, no. 5 (May 2007): 504–506. Accessed September 9, 2023. https://www.asprs.org/wp-content/uploads/pers/2007journal/may/lookback.pdf.

Colton, John. "Physics 123 Lecture 30 Warm-up Questions." BYU Physics and Astronomy, November 5, 2012. Accessed October 17, 2023. https://physics.byu.edu/faculty/colton/docs/phy123-fall12/jitt30a.html.

Lê, Hoàng-Ân. "Camera Model: Intrinsic Parameters." July 30, 2018. Accessed October 14, 2023. https://lhoangan.github.io/camera-params/.

Zhang, Zhengyou. "Camera Calibration," May 2007. Accessed October 10, 2023. https://people.cs.rutgers.edu/elgammal/classes/cs534/lectures/CameraCalibration-book-chapter.pdf.

# Appendix A Eigenvalue and Eigenvector Problem

# Appendix B Tait-Bryan Angles

$$R = R_z(\gamma)R_y(\beta)R_x(\alpha) \tag{B.1}$$

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix} \tag{B.2}$$

$$R_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & -\sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & -\cos(\beta) \end{bmatrix} \tag{B.3}$$

$$R_z(\gamma) = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{B.4}$$