IB Math Analysis and Approaches (HL)

Extended Essay

May 2024 Session

Understanding and Applying the Fundamental Principles of Camera Calibration

---

Research Question: What mathematical techniques can be employed to create precise camera models for accurately converting 2D image data into 3D point reconstructions, and what are their practical applications in real-world scenarios?

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.

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 to be used in applications requiring high accuracy. Combined with the potential for manufacturing defects as well as lens distortion effects 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.

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

The task of generating an accurate 3D model from various 2D images is a multi-step process, and consists of:

and they mainly fall into two main categories. The first method is based on the cross-referencing of keypoints between the images. Computer algorithms such as SIFT (Scale Invariant Feature Transform) and SURF (Speeded-Up Robust Features)

2 Tsai Camera Model

A camera model is a projection model which describes a mathematical relationship between points in 3D space and its projection onto an image plane. There are many variation of a

The Tsai camera model builds upon the pinhole camera model, which is one of the simplest and most commonly used camera models.

2.1 Pinhole Camera

A pinhole camera is a simple camera without a lens. Instead, it has a small aperture, and light rays pass through the aperture and projects an inverted image

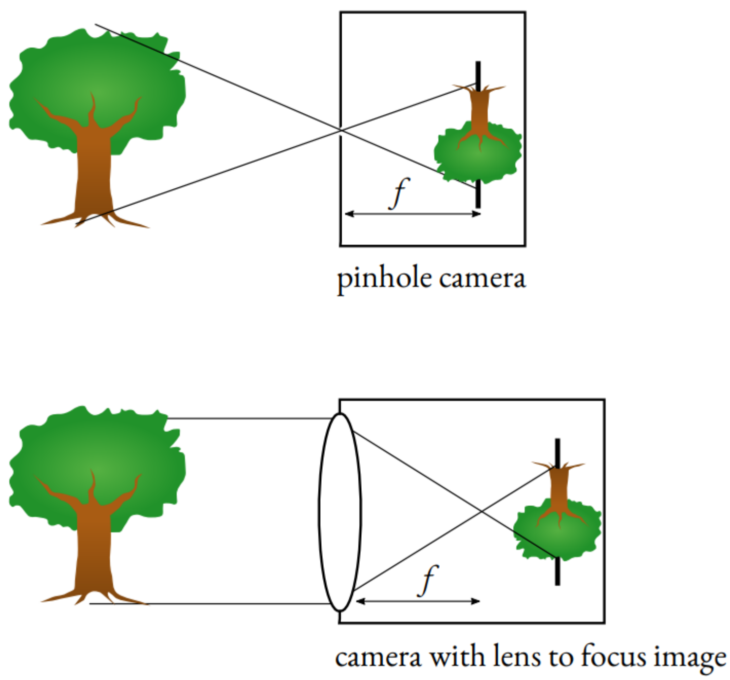


Figure 1: Difference between a pinhole camera and a lens camera

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

- 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,

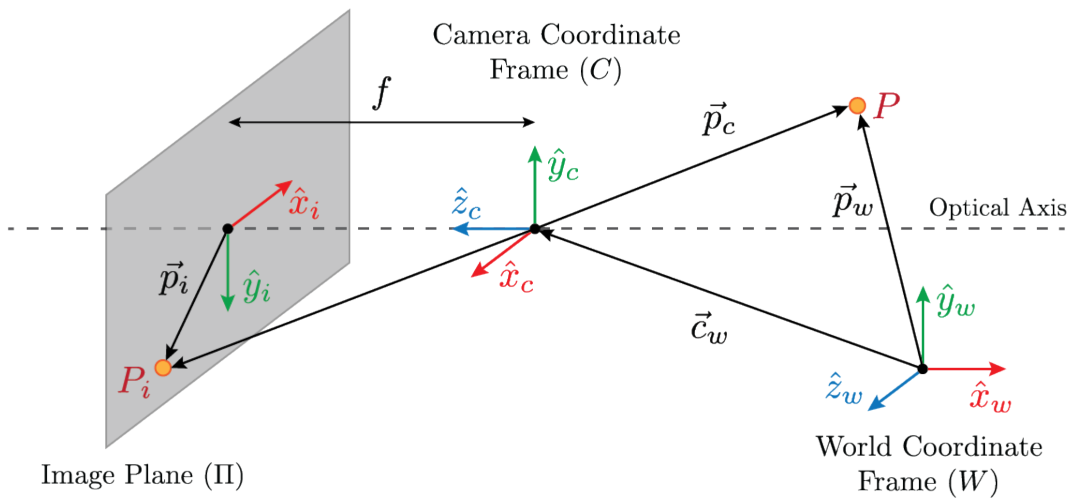
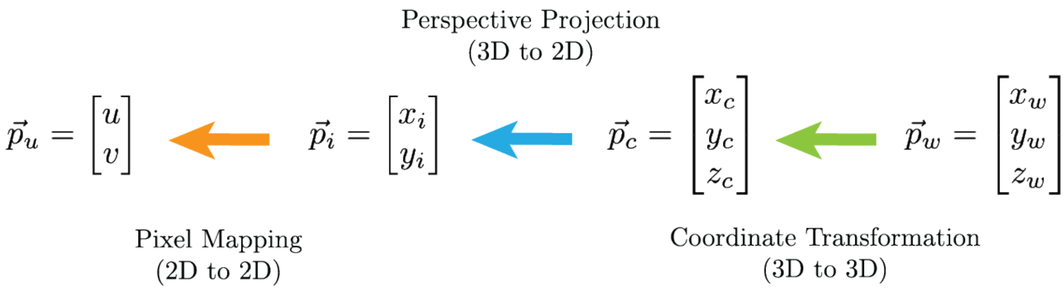


Figure 2: Pinhole Camera Model

2.2 Nomenclature

3 Projection Matrix



3.1 Intrinsic Parameters

$$\vec{p}_i = M_{int}\vec{p}_c \tag{1}$$

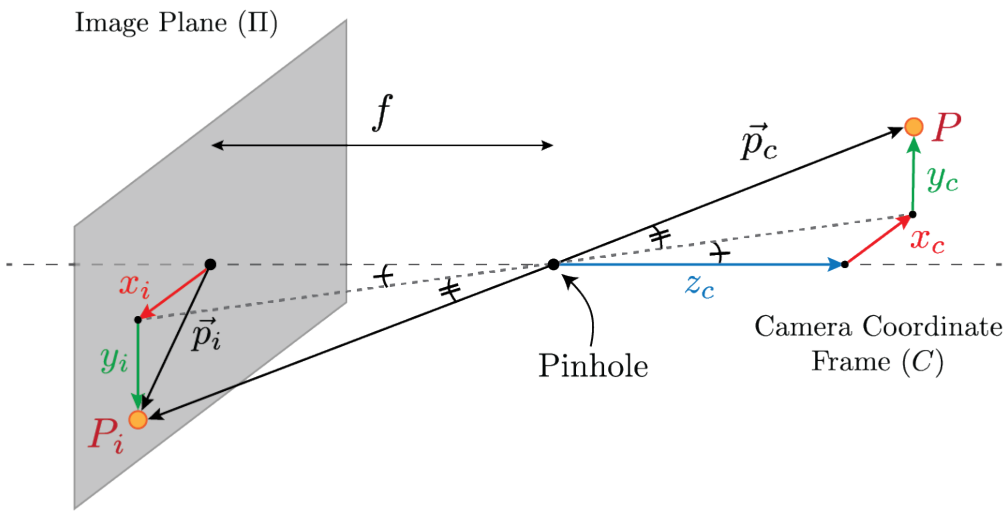
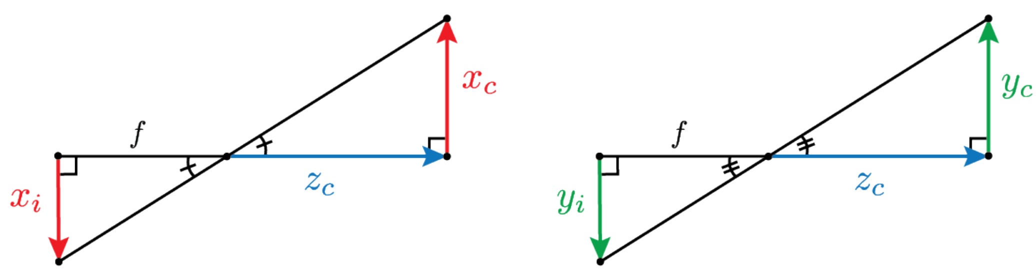


Figure 3: Perspective Projection



$$\frac{x_i}{f} = \frac{x_c}{z_c} \Rightarrow x_i = f\frac{x_c}{z_c} \tag{2}$$

$$\frac{y_i}{f} = \frac{y_c}{z_c} \Rightarrow y_i = f\frac{y_c}{z_c} \tag{3}$$

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.

$$u = m_x x_i + c_x \tag{4}$$

$$v = m_y y_i + c_y \tag{5}$$

$$u = m_x f \frac{x_c}{z_c} + c_x \tag{6}$$

$$v = m_y f \frac{y_c}{z_c} + c_y \tag{7}$$

$$u = f_x \frac{x_c}{z_c} + c_x \tag{8}$$

$$v = f_y \frac{y_c}{z_c} + c_y \tag{9}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong \begin{bmatrix} z_c u \\ z_c v \\ z_c \end{bmatrix} = \begin{bmatrix} f_x x_c + z_c c_x \\ f_y y_c + z_c c_y \\ z_c \end{bmatrix} = \underbrace{\begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{=M_{int}} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \tag{10}$$

3.2 Extrinsic Parameters

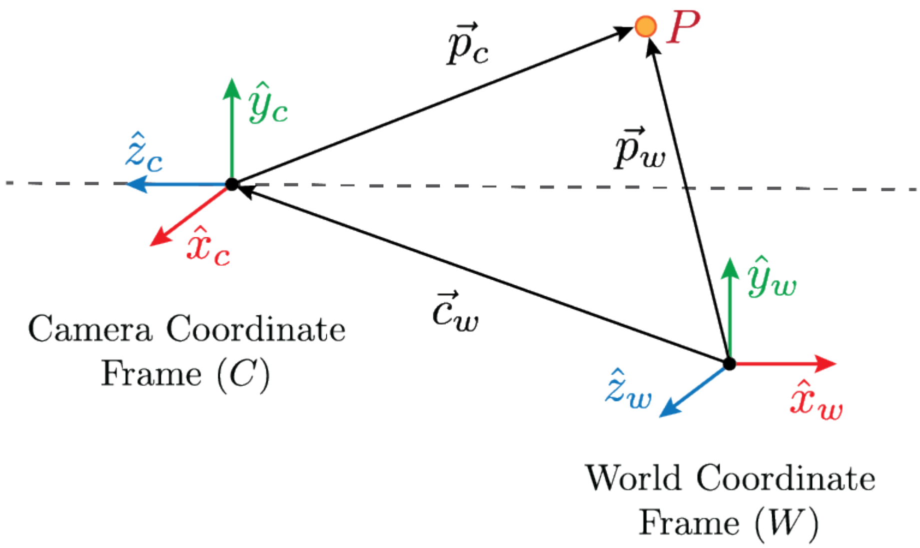


Figure 4: Coordinate Transformation

$$\vec{p}_c = M_{ext}\vec{p}_w \tag{11}$$

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

$$\vec{p}_c = R(\vec{p}_w - \vec{c}_w) \tag{13a}$$

$$= R\vec{p}_w - R\vec{c}_w \tag{13b}$$

$$\vec{p}_c = R\vec{p}_w + \vec{t} \tag{14}$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{15}$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{=M_{ext}} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{16}$$

3.3 Putting It Together

$$\vec{p}_{i_{(n)}} = \underbrace{M_{int}M_{ext}}_{=P}\vec{p}_{w_{(n)}} \tag{17}$$

$$\begin{bmatrix} u_n \\ v_n \\ 1 \end{bmatrix} \cong \begin{bmatrix} \widetilde{w}_n u_n \\ \widetilde{w}_n v_n \\ \widetilde{w}_n \end{bmatrix} = \begin{bmatrix} \widetilde{u}_n \\ \widetilde{v}_n \\ \widetilde{w}_n \end{bmatrix} = \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{=P} \begin{bmatrix} x_w^{(n)} \\ y_w^{(n)} \\ z_w^{(n)} \\ 1 \end{bmatrix} \tag{18}$$

$$\widetilde{u}_n = p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14} \tag{19}$$

$$\widetilde{v}_n = p_{21}x_w^{(n)} + p_{22}y_w^{(n)} + p_{23}z_w^{(n)} + p_{24} \tag{20}$$

$$\widetilde{w}_n = p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34} \tag{21}$$

$$u_n = \frac{\widetilde{u}_n}{\widetilde{w}_n} = \frac{p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14}}{p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}} \tag{22}$$

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

$$u_n(p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}) = p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14} \tag{24}$$

$$v_n(p_{31}x_w^{(n)} + p_{32}y_w^{(n)} + p_{33}z_w^{(n)} + p_{34}) = p_{21}x_w^{(n)} + p_{22}y_w^{(n)} + p_{23}z_w^{(n)} + p_{24} \tag{25}$$

$$0 = p_{11}x_w^{(n)} + p_{12}y_w^{(n)} + p_{13}z_w^{(n)} + p_{14} - p_{31}u_nx_w^{(n)} - p_{32}u_ny_w^{(n)} - p_{33}u_nz_w^{(n)} - p_{34}u_n \tag{26}$$

$$0 = p_{21}x_w^{(n)} + p_{22}y_w^{(n)} + p_{23}z_w^{(n)} + p_{24} - p_{31}v_nx_w^{(n)} - p_{32}v_ny_w^{(n)} - p_{33}v_nz_w^{(n)} - p_{34}v_n \tag{27}$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} x_w^{(1)} & y_w^{(1)} & z_w^{(1)} & 1 & 0 & 0 & 0 & 0 & -u_n x_w^{(1)} & -u_n y_w^{(1)} & -u_n z_w^{(1)} & -u_1 \\ 0 & 0 & 0 & 0 & x_w^{(1)} & y_w^{(1)} & z_w^{(1)} & 1 & -v_n x_w^{(1)} & -v_n y_w^{(1)} & -v_n z_w^{(1)} & -v_1 \\ x_w^{(2)} & y_w^{(2)} & z_w^{(2)} & 1 & 0 & 0 & 0 & 0 & -u_n x_w^{(2)} & -u_n y_w^{(2)} & -u_n z_w^{(2)} & -u_2 \\ 0 & 0 & 0 & 0 & x_w^{(2)} & y_w^{(2)} & z_w^{(2)} & 1 & -v_n x_w^{(2)} & -v_n y_w^{(2)} & -v_n z_w^{(2)} & -v_2 \\ & \vdots & & & & & \vdots & & & \vdots & & \\ x_w^{(n)} & y_w^{(n)} & z_w^{(n)} & 1 & 0 & 0 & 0 & 0 & -u_n x_w^{(n)} & -u_n y_w^{(n)} & -u_n z_w^{(n)} & -u_n \\ 0 & 0 & 0 & 0 & x_w^{(n)} & y_w^{(n)} & z_w^{(n)} & 1 & -v_n x_w^{(n)} & -v_n y_w^{(n)} & -v_n z_w^{(n)} & -v_n \end{bmatrix}}_{=A} \underbrace{\begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix}}_{=p} \tag{28}$$

Thus, using

4 Geometric Distortion

4.1 Symmetrical Lens Distortion

4.2 Asymetrical Lens Distortion

4.3 De-centering (Tangential) Lens Distortion

5 Applications

Acknowledgements

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

References

[1] R. Sensing and P. Panel, *Photogrammetry and remote sensing*, 2017. [Online]. Available: https://web.archive.org/web/20170830062535/https://www.cices.org/pdf/P%26RSinformation.pdf.