Concept Review

# Camera Model

## Why Model a Camera?

Cameras are common place items today, often a must both from a technological perspective and user's point of view. Cell phones, cars, gaming consoles, drones etc. all deploy this sensor due to it's innate ability to provide vision or sight similar to human eyes. Modeling a camera allows us to understand the limitations of projecting a real 3D world onto an image plane, and how features identified in images could be localized and placed back in the real world. While quite a few types of camera sensors exist with corresponding models, this document will focus on the pinhold camera model.

## Camera Extrinsic Properties

A camera's extrinsic definition makes use of a homogeneous transformation matrix to describe a location $\boldsymbol{t_j^i}$ and rotation $\boldsymbol{R_j^i}$ with respect to a desired reference frame. For cameras the extrinsic definition is used to convert a point of light $\overrightarrow{\boldsymbol{P_E}}$ from an environment frame to the camera reference frame. The general form for a homogeneous transformation matrix is,

.

$$\boldsymbol{T_j^i} = \begin{bmatrix} \boldsymbol{R_j^i} & \boldsymbol{t_j^i} \\ \boldsymbol{0} & \boldsymbol{1} \end{bmatrix} \tag{1}$$

The rotation matrix $\boldsymbol{R_j^i}$ can be constructed as the combination of 3 successive rotation about the principal axes $x, y, z$ specified in a particular order. $\boldsymbol{R_j^i}$ defines the relation of frame $j$ with respect to frame $i$. Rotations about each principal axis: (Spong, 2006)

$$R_{x,\phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{bmatrix} \qquad R_{y,\theta} = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \qquad R_{z,\psi} = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation about $x$ axis | Rotation about $y$ axis | Rotation about $z$ axis

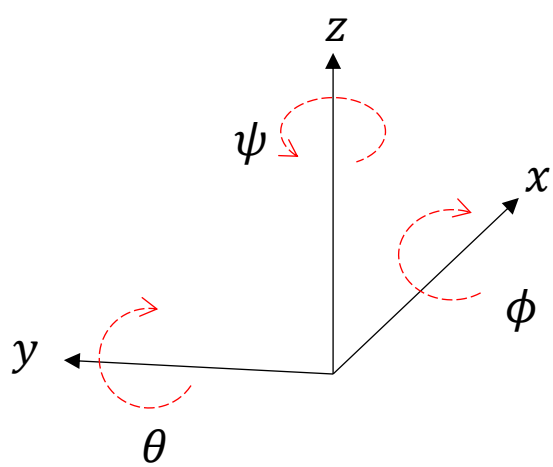


Figure 1: Positive rotation for primary x-y-z axis

With a rotation definition specified the conversion of a point of light $\overrightarrow{\boldsymbol{P_E}}$ from the environment frame $\boldsymbol{E}$ to the camera frame $\boldsymbol{C}$ is,

$$\overrightarrow{P_c} = T_E^c \overrightarrow{P_E} \tag{2}$$

Having defined the point of light in the camera frame the next step is to construct the conversion from a physical point to a pixel in an image.

## Pinhole Camera Model

The pinhole camera model is the first step for converting a physical point of light into an image coordinate. Lens distortion is considered later in the camera model formulation. Point of light $\overrightarrow{P_c}$ originates from an optical center called the optical focal point.

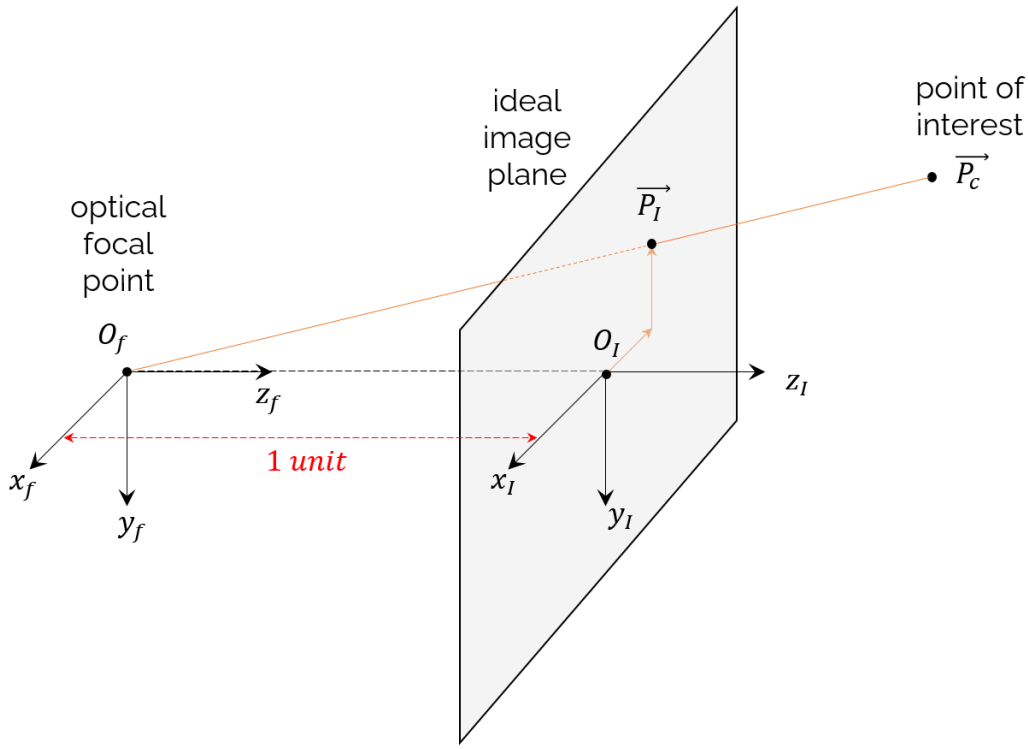


Figure 2: Description of Image Focal point and Ideal image plane. (Waslander, 2021)

Point $\overrightarrow{P_C}$ is used to describe the relationship between the optical axis $O_f x_f y_f z_f$ and the ideal image plane originating at $O_I x_I y_I z_I$. Using image projection and similarity triangles point $\overrightarrow{P_c}$ can be projected onto the ideal image plane using the following relationship,

$$\overrightarrow{P_I} = \begin{bmatrix} Px_I \\ Py_I \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{X}{Z} \\ \frac{Y}{Z} \\ 1 \end{bmatrix} \tag{3}$$

By projection point $\overrightarrow{P_c}$ to the ideal image frame there is a change from $\mathbb{R}^3$to $\mathbb{R}^2$ space. Now in the ideal image plane we consider image distortion due to the type of lens used. The two most common distortion components are,

- Radial distortion ($\boldsymbol{k_{rad}}$)
- Tangential distortion ($\boldsymbol{k_{tan}}$) .

Radial distortion originates at the center of the image plane and depending on the sign of the leading distortion coefficient, can cause an image to warp inwards in a barrel like fashion, or

outwards. The combination of both radial and tangential distortion on the ideal image point $\overrightarrow{P_I}$ can be written as,

$$\vec{P}_{Dist} = \begin{bmatrix} Px_{Dist} \\ Py_{Dist} \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4 + K_3 r^6) \begin{bmatrix} Px_I \\ Py_I \end{bmatrix} + \begin{bmatrix} 2\tau_1 Px_I Py_I + \tau_2 (r^2 + 2Px_I^2) \\ 2\tau_2 Px_I Py_I + \tau_1 (r^2 + 2Py_I^2 \end{bmatrix} \quad (4)$$

where $r = \sqrt{Px_I^2 + Py_I^2}$ (Barfoot, 2014). An examples of radial distortion is shown in Figure 4.

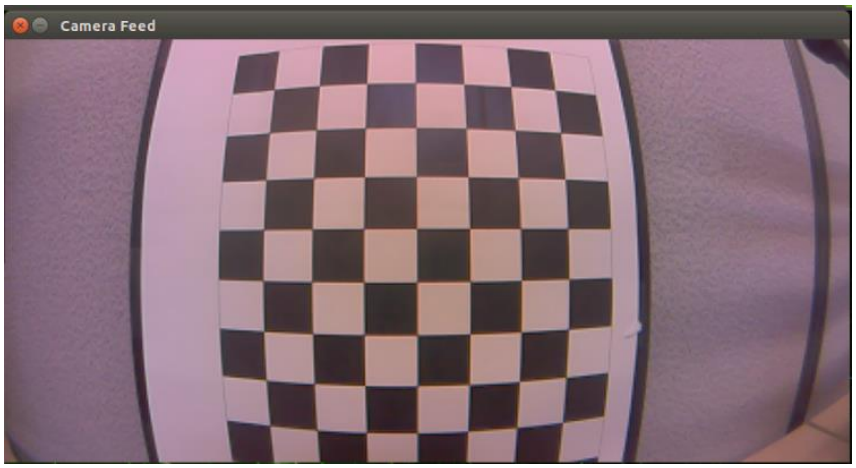


Figure 3: Camera image with barrel distortion.

With the ideal image points modified by the distortion model we now transform from physical coordinates in the ideal image plane to coordinates in the digital image frame. When describing the location of a pixel in an image the following convention is used,

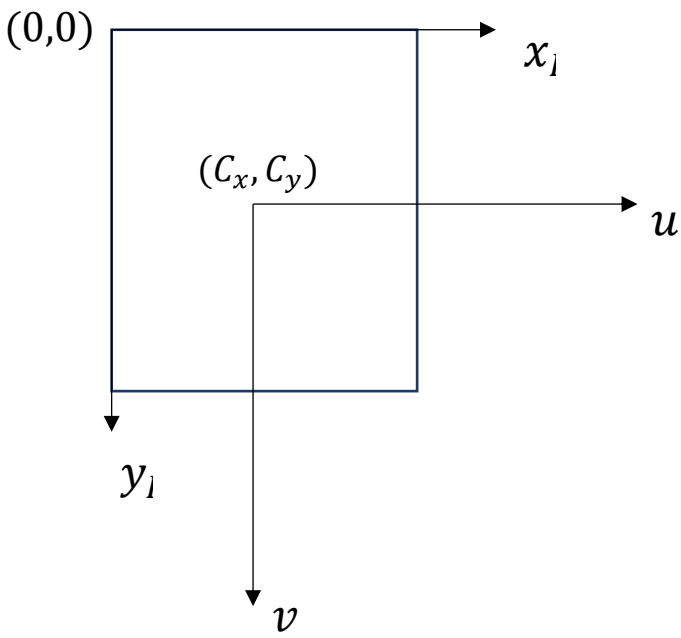


Figure 4: Description of Ideal image to digital image.

With the $u$ axis conciding with $x_I$ and $v$ axis coinciding with $y_I$ as seen in Figure 4. The transformation between the ideal image plane and the camera image plane is defined as,

$$\begin{matrix} u \\ v \end{matrix} = \begin{matrix} f_x Px_{Dist} + C_x \\ f_y Py_{Dist} + C_y \end{matrix} \tag{5}$$

Where $f_x$ and $f_y$ are the focal lengths along the $x$ and $y$ directions. $C_x$ and $C_y$ are centers of the image. All 4 of these parameters are dependent on image resolution. (Barfoot, 2014)

These steps take us from a 3D point to the image frame. Usually, we need to do the reverse and either calculate 3D information based on information from the camera or rectify distorted image data for a give image processing pipeline. A camera calibration process is used to calculate the camera matrix and distortion parameters.

## Summary

Cameras can be modeled using the Camera Obscura (Pinhole camera) model with the inclusion of distortion components. The sequence for transforming a point $\overrightarrow{P_E}$ from the global frame to a pixel coordinate $(u, v)$ is,

Camera Frame Transformation

$$\overrightarrow{P_C} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0 & 1 \end{bmatrix} \overrightarrow{P_E}$$

Image Plane projection

$$\overrightarrow{P_I} = \begin{bmatrix} Px_I \\ Py_I \\ 1 \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \\ Z/Z \end{bmatrix} = \frac{\overrightarrow{P_c}}{Z}$$

Image distortion

$$\vec{P}_{Dist} = \begin{bmatrix} Px_{Dist} \\ Py_{Dist} \\ 1 \end{bmatrix} = f(\boldsymbol{k_{rad}}, \boldsymbol{k_{tan}}, \overrightarrow{P_I})$$

Pixel Image coordinates

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & C_x \\ 0 & f_y & C_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Px_{Dist} \\ Py_{Dist} \\ 1 \end{bmatrix}$$

© 2022 Quanser Inc., All rights reserved.

Quanser Inc.
119 Spy Court
Markham, Ontario
L3R 5H6
Canada

info@quanser.com
Phone: 19059403575
Fax: 19059403576
Printed in Markham, Ontario.

For more information on the solutions Quanser Inc. offers, please visit the web site at:
http://www.quanser.com

This document and the software described in it are provided subject to a license agreement. Neither the software nor this document may be used or copied except as specified under the terms of that license agreement. Quanser Inc. grants the following rights: a) The right to reproduce the work, to incorporate the work into one or more collections, and to reproduce the work as incorporated in the collections, b) to create and reproduce adaptations provided reasonable steps are taken to clearly identify the changes that were made to the original work, c) to distribute and publicly perform the work including as incorporated in collections, and d) to distribute and publicly perform adaptations. The above rights may be exercised in all media and formats whether now known or hereafter devised. These rights are granted subject to and limited by the following restrictions: a) You may not exercise any of the rights granted to You in above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation, and b) You must keep intact all copyright notices for the Work and provide the name Quanser Inc. for attribution. These restrictions may not be waved without express prior written permission of Quanser Inc.