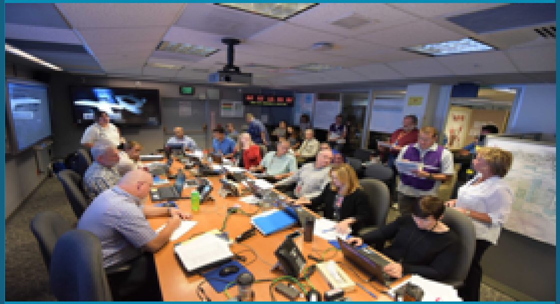
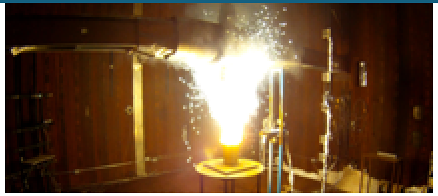
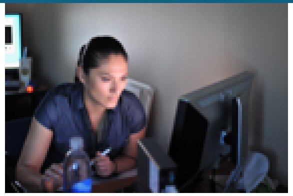
SAND2022-1305C

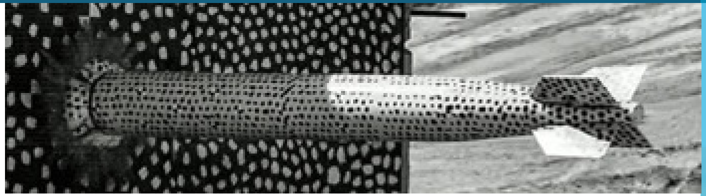
Sandia National Laboratories

Imaging Basics









IMAC XL Short Course: Optical Techniques for Experimental Modal Analysis

Dan Rohe, Bryan Witt, and Phil Reu

U.S. DEPARTMENT OF ENERGY NNSA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

General Camera Concepts and Terminology

Camera Geometry and the Pinhole Camera Model

Reconstructing 3D Positions from 2D positions on the Image

Camera Calibration

Camera Selection for a Given Test

Review

General Camera Concepts and Terminology

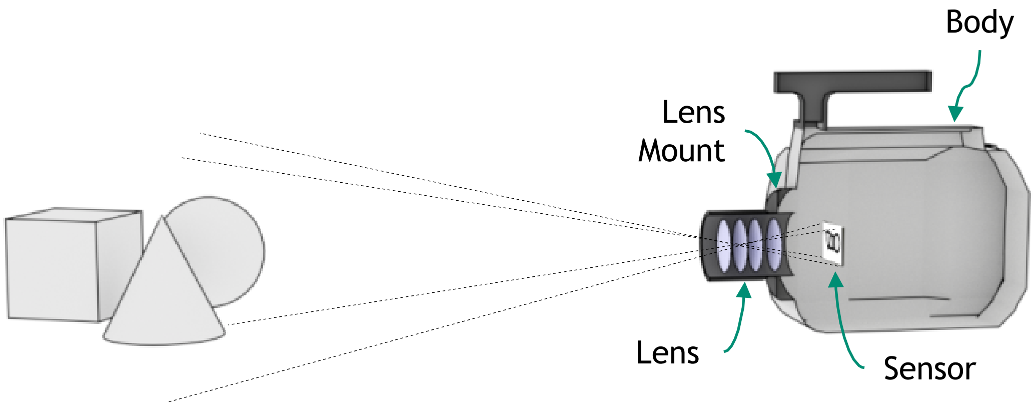
Camera Components and Terminology

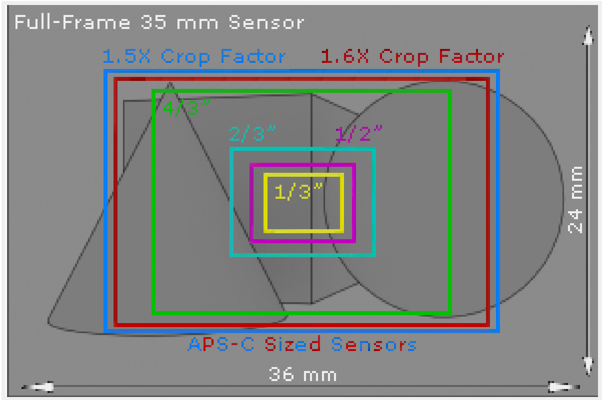
Body/Housing – The general structure of the camera that you can hold on to or mount to
- We require our cameras to be held very still during optical testing

Lens Mount – The portion of a camera to which a lens gets mounted
- Certain lenses require certain mounts
- Lenses with autofocus or automatic aperture may require electronic interfaces in the lens mount

Lens – The portion of the camera through which light refracts and is focused

Sensor – The portion of the camera that records the image
- Made up of perhaps millions of individual light sensors, with one (for monochrome) or more (for color images) sensors per pixel on the image
- Pixel size is often computed via sensor size divided by number of pixels
- Sensor size is often described via a "crop factor" relative to a full frame (35mm film) sensor.
- A smaller sensor will tend to effectively "zoom in" a given image, similarly to cropping off the edges of an image





Lens Components and Terminology

Focal Length – Distance between the pinhole and the image plane
- Together with sensor size, determines view angle
- Lenses are often described with a 35mm or “full frame” equivalent focal length. The “effective focal length” would be the crop factor times the focal length.

Aperture – The size of the opening in the lens
- Larger apertures let through more light at the expense of reduced depth-of-field
- Many lenses have adjustable apertures to adjust exposure
- Described relative to the focal length (f/5.6, f/11)
- Because the so-called f-number is in the denominator, smaller numbers correspond to larger apertures
- Many consumer lenses support setting the aperture electronically, ensure the camera body and lens mount you are using supports electronic aperture if the lens doesn’t have a manual aperture option.

Maximum Aperture – The largest aperture supported by a lens
- Lenses that have larger maximum apertures are often called “fast” lenses because they can get an equivalent exposure with a shorter exposure time
- Larger aperture lenses are often more expensive, requiring larger or more complex optics and higher quality glass

Auto-Focus – The ability of a lens to automatically focus on an object
- Many consumer lenses support auto-focus, ensure the camera body and lens mount you are using supports auto-focus if the lens doesn’t have a manual focus option.

Decoding Lens Parameters:

Nikon Nikkor 18-140 mm 1:3.5-5.6

Zoom Lens with focal lengths between 18 and 140 mm. Maximum aperture is f/3.5 at widest angle and f/5.6 at maximum zoom.

Lens manufacturers may have various other items in name:
- Lens Model
- Autofocus
- Image stabilization
- Lens generation
- Lens Mount Type
- Full Frame vs Crop Sensor

Lens Components and Terminology, cont'd

Minimum Focus Distance – The closest distance a camera can be to an object while still acquiring an in-focus image.

- Macro lenses will typically have the ability to focus very close to the lens, allowing users to achieve a large magnification of the object on the sensor

Magnification – The scale of the object on the camera sensor compared to the scale of the object in real life

- Depends on the minimum focus distance and the focal length of the lens
- Macro lenses will typically have magnifications such as 2:1, 1:1, or 1:2

Depth of Field – The amount of the field of view that is in focus at one time

- If the depth of field is large, the camera can image objects close to the camera and far away from the camera simultaneously
- If the depth of field is small, objects far from the camera will be blurry when the camera is focused up close, and vice versa

Distortions – Imperfections in the lens or camera optics

- Many types of distortions exist for cameras and lenses

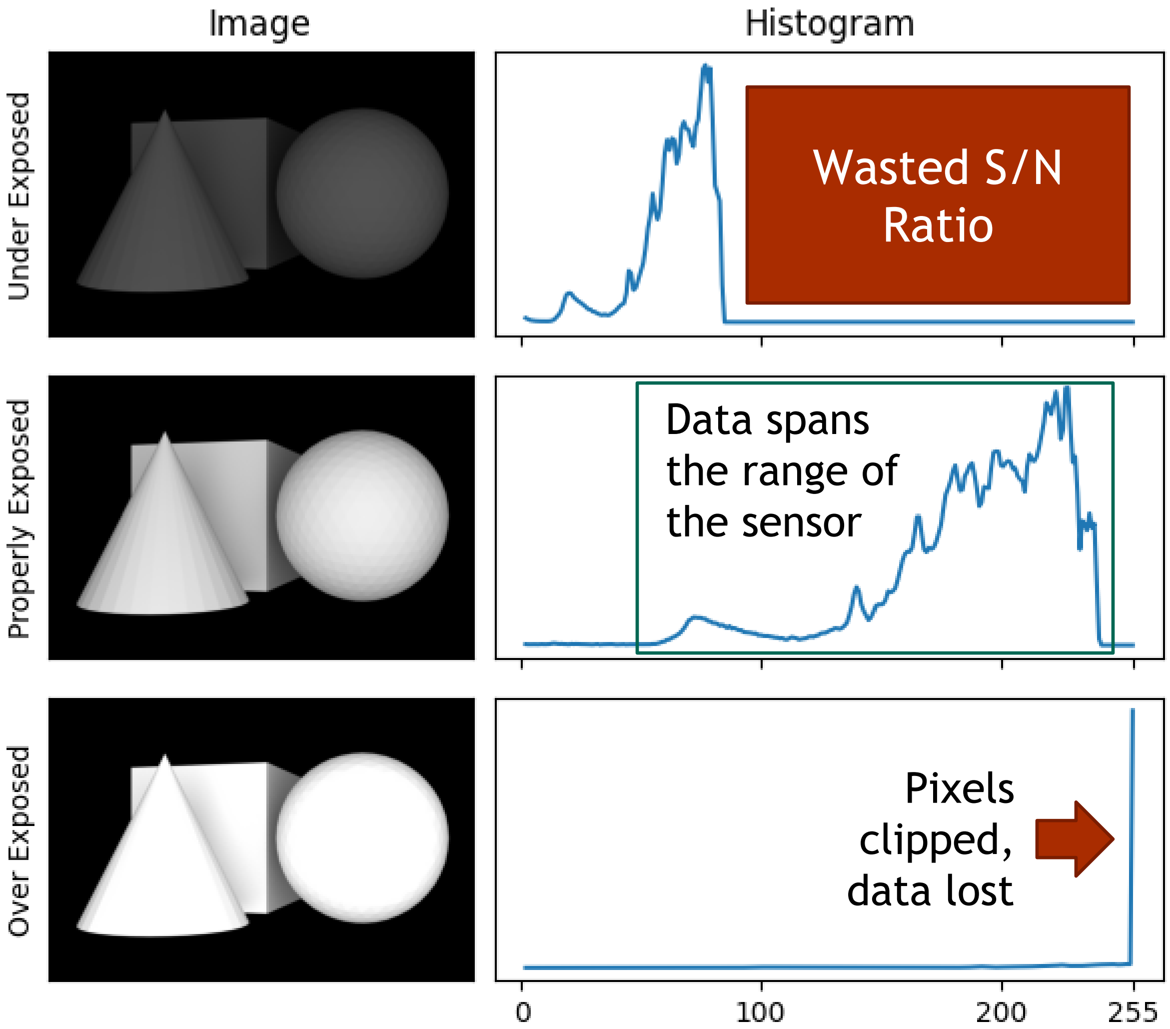
Exposure Concepts and Terminology

Exposure – The general brightness or darkness of the image

- Under-exposed images result in noisy data
- Over-exposed images result in clipping and data loss
- Use the histogram to analyze exposure of an image.

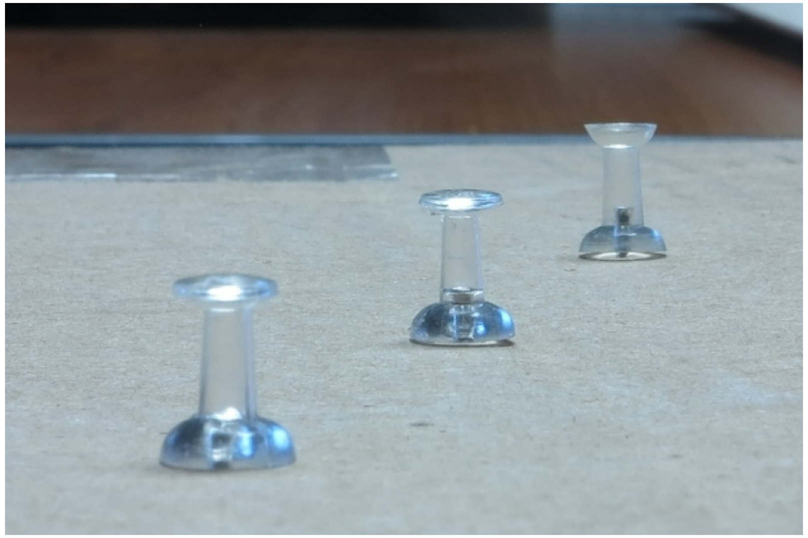
Exposure is determined by:

- Shutter/Exposure Time – The time the camera is collecting light
  - Increase in shutter time can result in motion blur if the camera or part moves significantly during the time the shutter is open
  - Decreasing the shutter time will tend to freeze motion better
- Aperture Size – More light is collected per time if the aperture is larger
  - Increasing the aperture will result in narrower depth of field, causing objects that are not at the focus distance to blur
  - Decreasing the aperture will result in larger depth of field, allowing more of the image to be in focus
  - Decreasing aperture too much will also introduce diffraction effects
- Gain/ISO – Scaling applied to the image
  - Scaling the brightness of the image also scales noise in the image, so generally better to adjust exposure via aperture or shutter time
  - Note that Gain or ISO is typically applied to the analog sensor before digitization. Therefore, it is typically better to apply gain or adjust ISO on the camera than to acquire underexposed images and then brighten with post-processing.
- Lighting – Just add more light to the scene!
  - Be aware that lights get hot, and can introduce heat waves between the test article and the camera

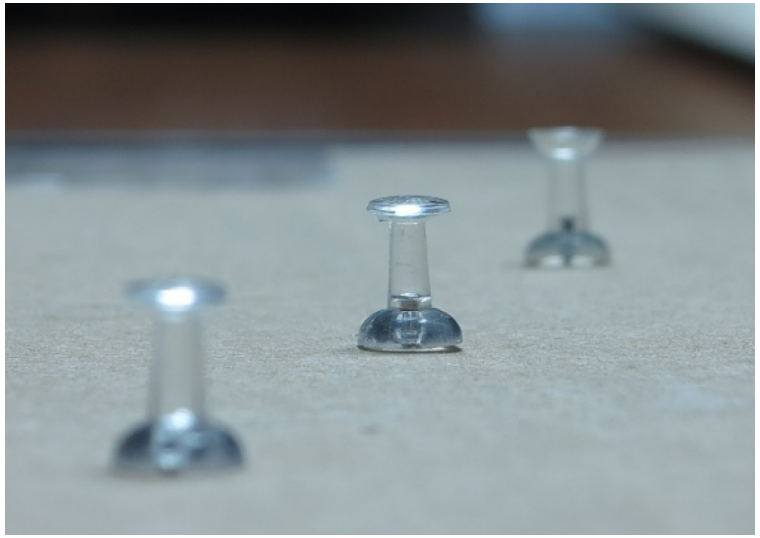


Exposure Example

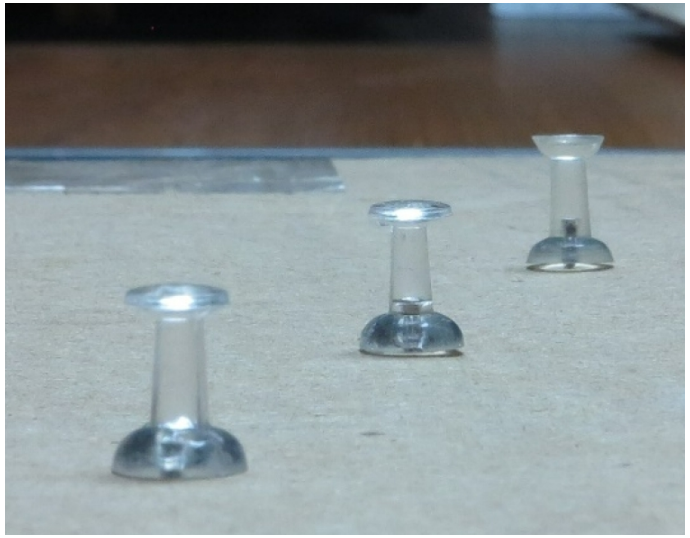
Three images with similar exposures using different techniques.



Exposure: 2 Second
Aperture: f/16
ISO: 100



Exposure: 1/6 second
Aperture: f/4.5
ISO: 100



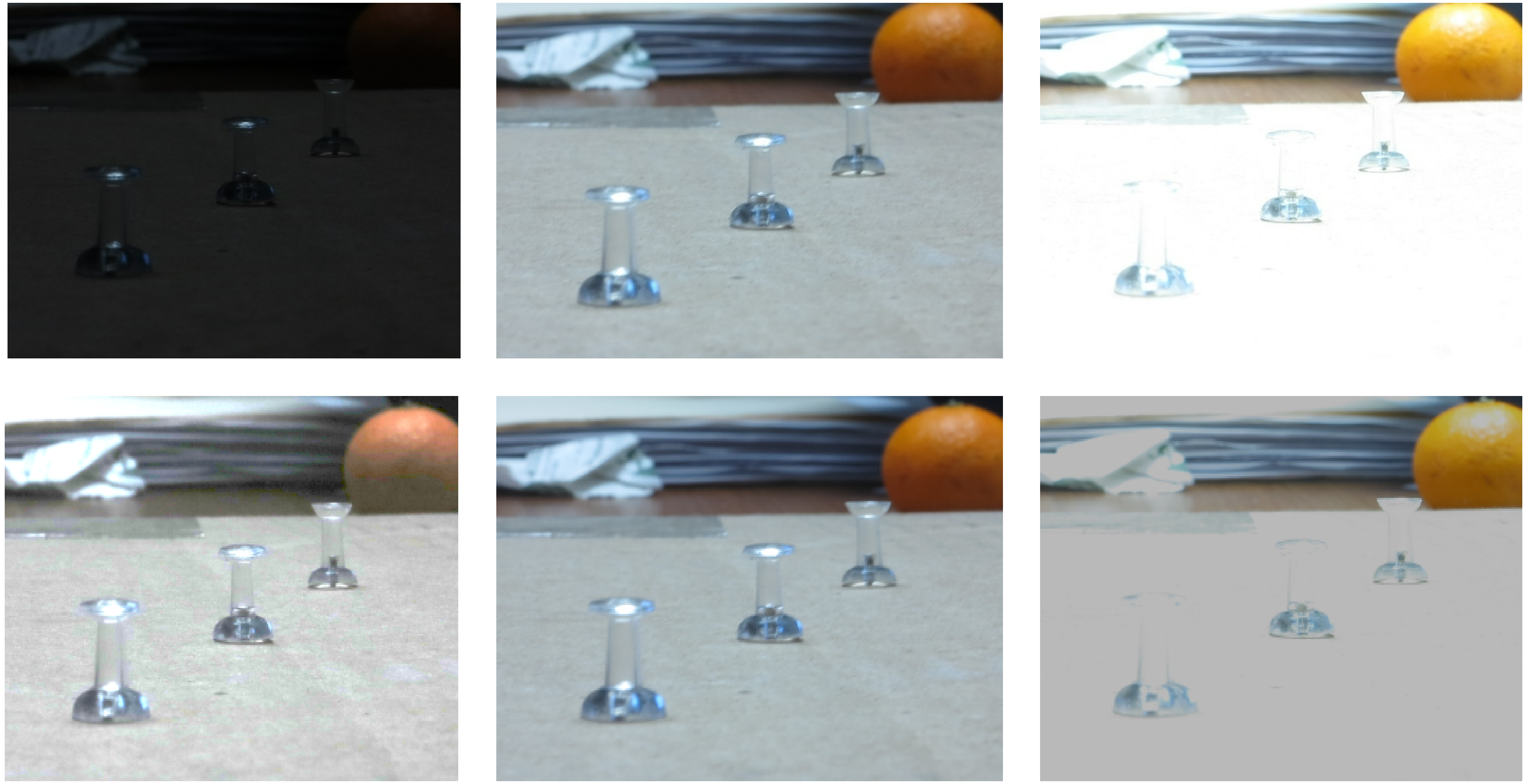
Exposure: 1/6 Second
Aperture: f/16
ISO: 1250

Over- and Under-Exposure Example

Images taken with varying exposures then corrected to have similar exposures

Camera Image

Corrected Image



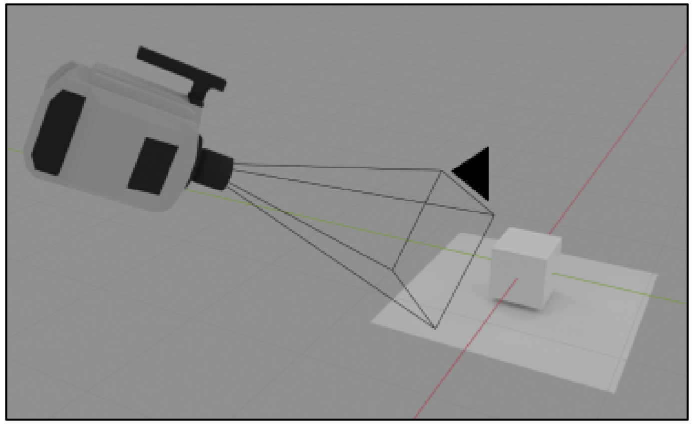
Underexposure results in poor signal-to-noise ratio	Proper exposure results in best signal-to-noise ratio	Over-exposure results in data lost, cannot recover

Camera Geometry and the Pinhole Model

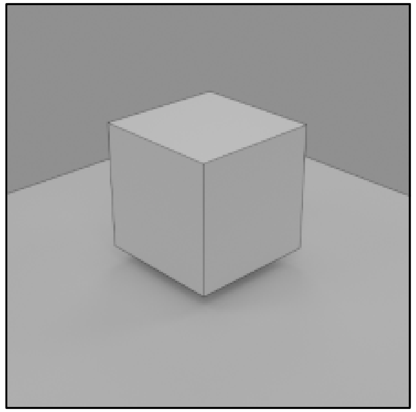
Mathematical Models for Cameras

A camera maps 3D positions in the world to 2D positions on an image

Cameras can be idealized using the **pinhole camera model** that projects 3D space to 2D positions on an image



3D Scene Imaged by a Camera

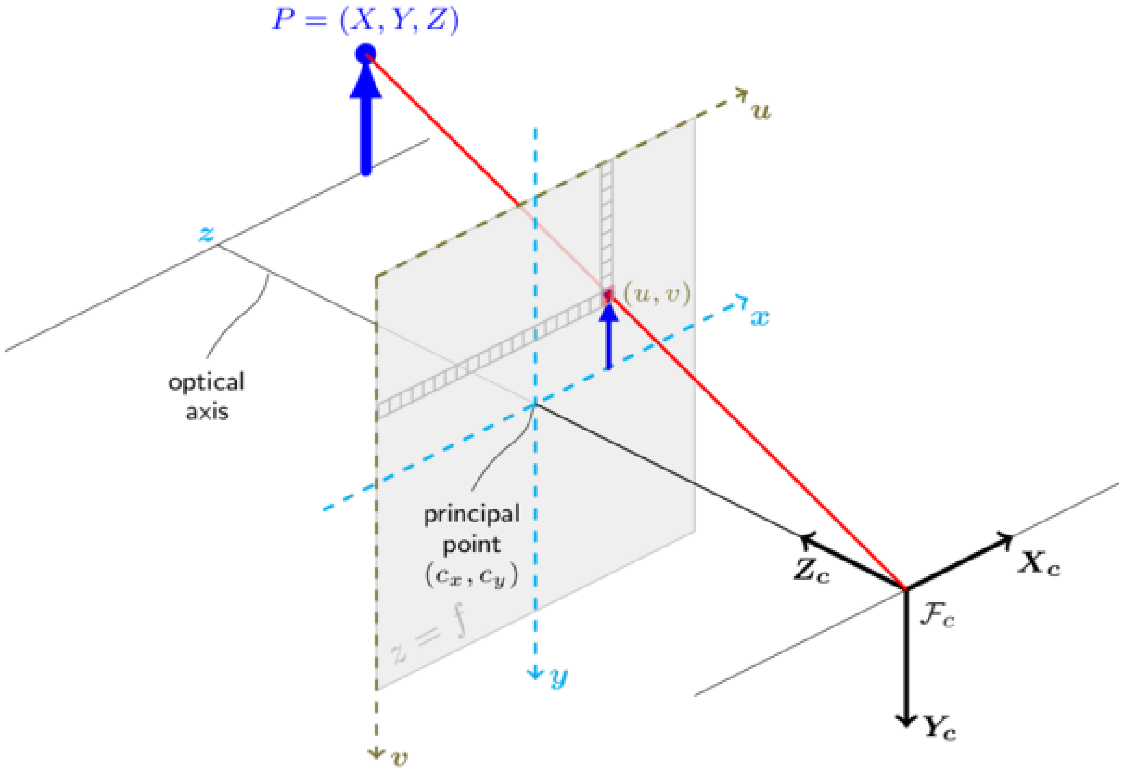


2D Image created by a Camera

Mapping 3D points to 2D with the Pinhole Camera Model

Given a camera at the origin of the c coordinate system pointed along the Z_c axis:

- Project the point $P = (X, Y, Z)$ onto the plane where $Z = f$
- Using similar triangles, we can find the intersection of the vector to point P with the plane $Z = f$
  - $X' = \frac{fX}{Z}$
  - $Y' = \frac{fY}{Z}$
- The final point on the image (u, v) can be recovered by:
  - $u = X' + c_x$
  - $v = Y' + c_y$
- The camera equations are thus
  - $u = \frac{fX}{Z} + c_x$
  - $v = \frac{fY}{Z} + c_y$
- The quantity f can be approximated by $f = f_{mm}/p_{mm}$ where f_{mm} is the focal length of the lens in mm, and p_{mm} is the size of a pixel on the sensor in mm.
  - Size of a pixel can be approximated as image resolution divided by sensor size.



Pinhole Camera Model, shown graphically [1]

[1] https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html

Aside: Homogeneous Coordinates

Commonly, we represent coordinates in 3D space and 2D space with 3-vectors and 2-vectors, but sometimes this is insufficient.

Examine the coordinate transformation:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

where the coordinate vector $[X\ Y\ Z]^T$ is modified by a rotation matrix and a translation vector to form vector $[X'\ Y'\ Z']^T$.

It would be very useful to represent this expression as one matrix multiplication rather than a matrix multiplication and an addition.

By adding a fourth coordinate to the coordinate vectors, we transform to homogeneous coordinates which enables that operation to be performed.

$$c \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} & T_x \\ R_{yx} & R_{yy} & R_{yz} & T_y \\ R_{zx} & R_{zy} & R_{zz} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Noting the free parameter c is used to normalize the resulting equations. To recover $[X'\ Y'\ Z']^T$, the vector $[X'\ Y'\ Z'\ 1]^T$ should be normalized so that the fourth row is 1.

- Note, the free parameter c is often left out of the equations, and assumed to be there. The assumption is always that you must scale that last row to 1 to achieve the correct parameters in the other rows.

This formulation allows for many more operations, including translations and projections, to be treated as linear operations.

The Pinhole Camera Model in Homogeneous Coordinates

We can transform the pinhole equations

- $u = \frac{fX}{Z} + c_x$
- $v = \frac{fY}{Z} + c_y$

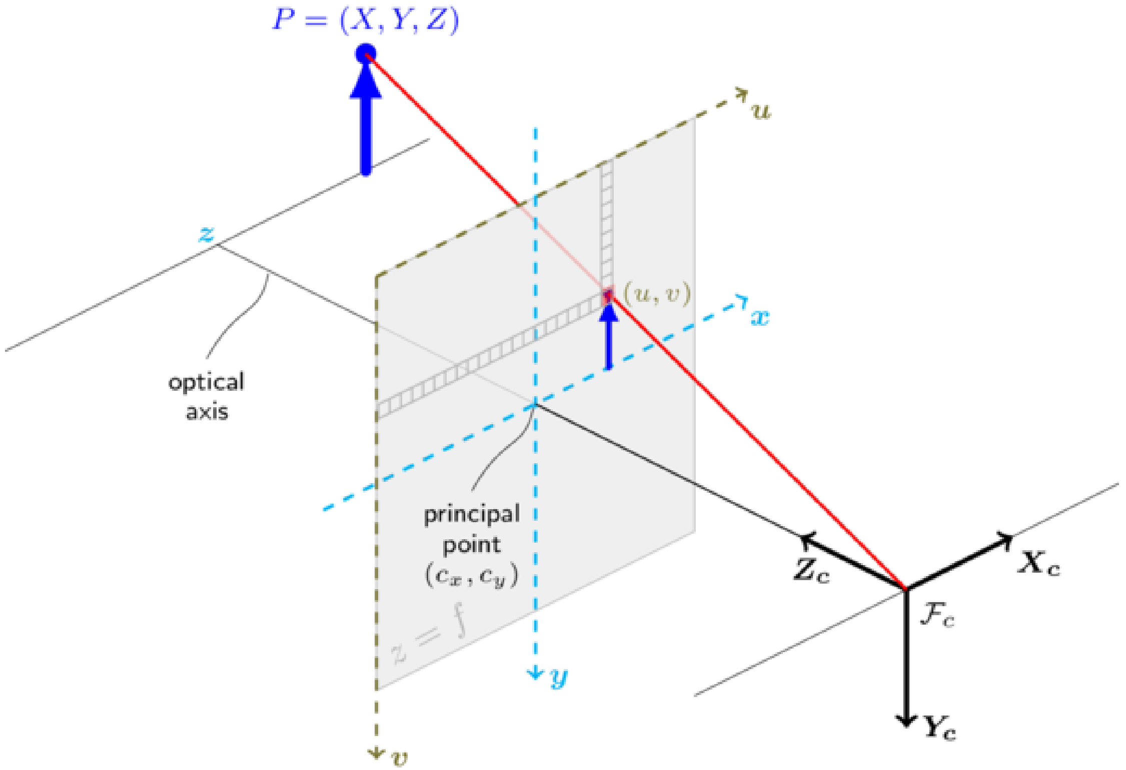
into homogeneous coordinates

- $c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

which is equivalent to

- $cu = fX + Zc_x$
- $cv = fY + Zc_y$
- $c = Z$

Dividing each expression by c recovers the original equations



Pinhole Camera Model, shown graphically [1]

[1] https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html

Extending the Pinhole Camera Model

The Pinhole camera model shown previously has some assumptions that are not necessarily accurate for real cameras.

- E.g. if the sensor is mounted slightly askew, it may not be perfectly perpendicular to the imaging plane of the lens.

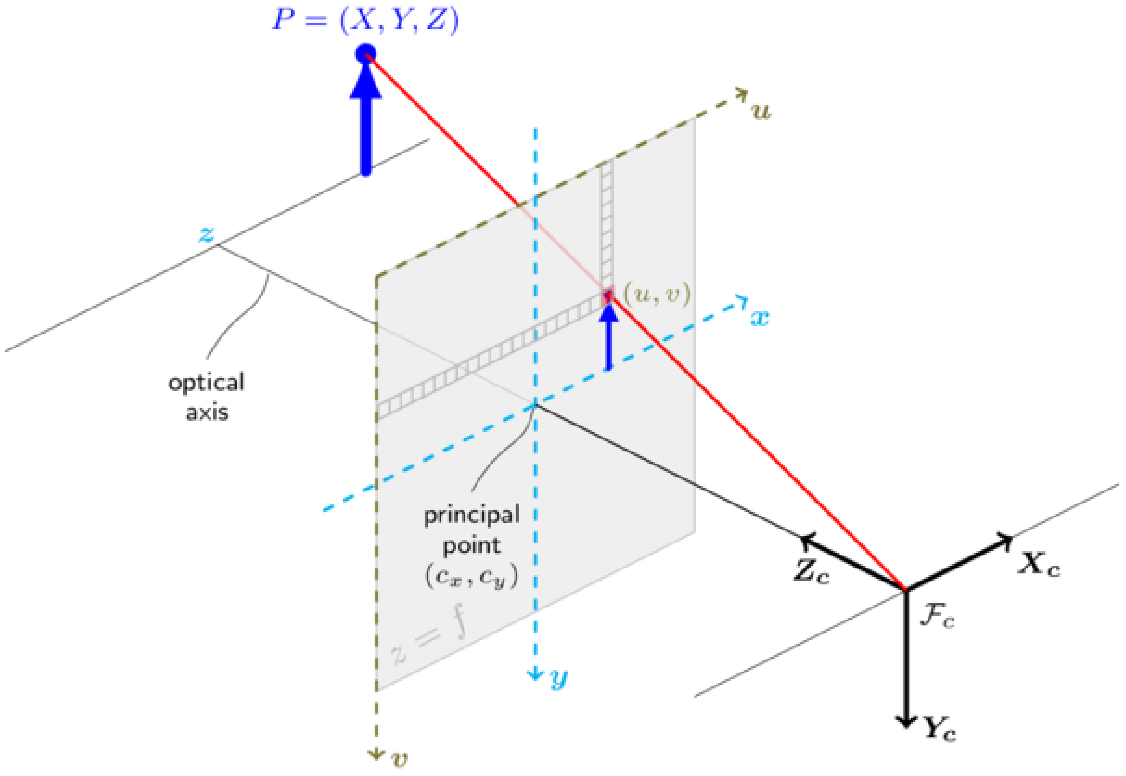
A more general pinhole camera equation that allows for misalignments in the sensor is

$$c\begin{bmatrix}u\\v\\1\end{bmatrix} = \begin{bmatrix}f_x & s & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}X\\Y\\Z\end{bmatrix}$$

where the focal distance can vary in the horizontal and vertical directions $f_x \neq f_y$ and a coupling between directions can also occur $s \neq 0$

We will define the above matrix **K**, the **intrinsic** camera matrix

$$\mathbf{K} = \begin{bmatrix}f_x & s & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{bmatrix}$$



Pinhole Camera Model, shown graphically [1]

Camera Located Somewhere in Space

Cameras will not generally be located at the origin and oriented along the principal axes.

We will therefore append the intrinsic camera matrix with an **extrinsic** camera matrix, which includes a rotation matrix $\mathbf{R}$ and translation vector $\mathbf{T}$

$$[\mathbf{R}|\mathbf{T}] = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} & T_x \\ R_{yx} & R_{yy} & R_{yz} & T_y \\ R_{zx} & R_{zy} & R_{zz} & T_z \end{bmatrix}$$

The final pinhole camera equations that map points in 3D space to points in the image is:

$$c\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} & T_x \\ R_{yx} & R_{yy} & R_{yz} & T_y \\ R_{zx} & R_{zy} & R_{zz} & T_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{R}|\mathbf{T}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Where the matrix $\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{T}]$ is often called the camera projection matrix.

Note: The vector $\mathbf{T}$ specifies the location of the world coordinate system origin in the camera coordinate system. The location of the camera pinhole in the world coordinate system can be recovered by $-\mathbf{R}^T\mathbf{T}$

Adding Lens Distortions

In a pinhole camera, straight lines in the world coordinate system are transformed into straight lines on the image.

Real lenses generally include distortions that transform straight lines into curved lines.

Distortions are generally modelled as functions of the ideal pinhole position on the image

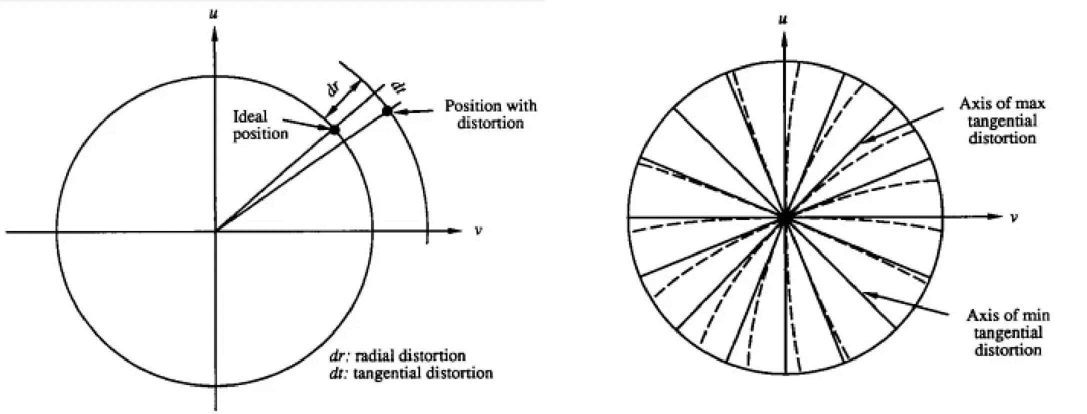
$$u_d = f_{du}(u, v)$$
$$v_d = f_{dv}(u, v)$$

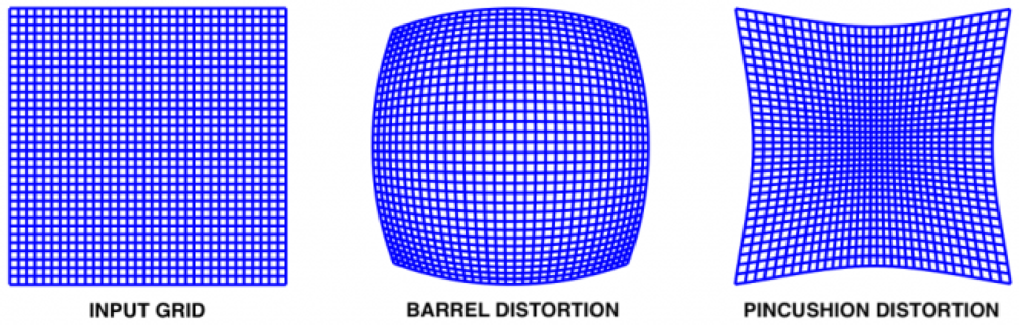
Depending on the software used to perform the camera calibration, these distortion models will vary.

- E.g in OpenCV

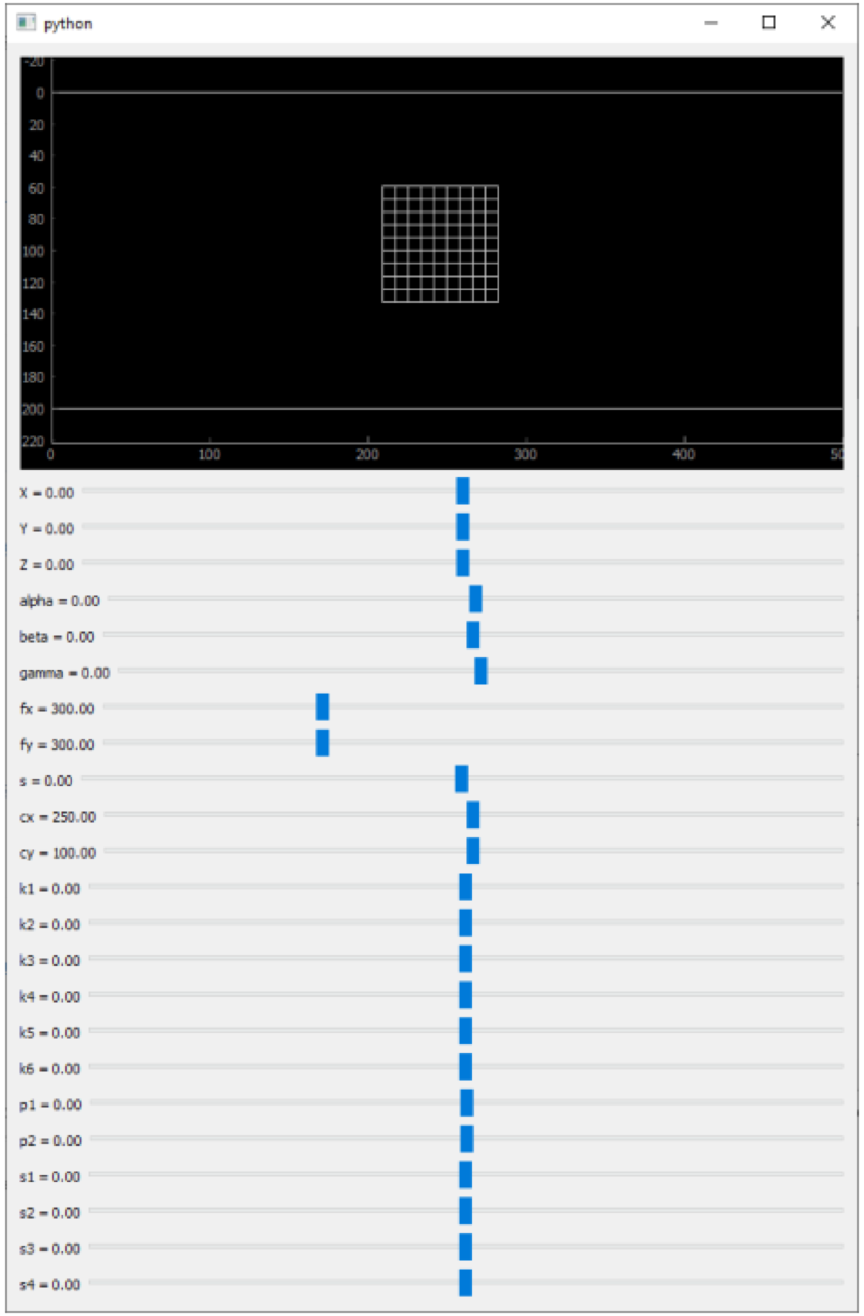
$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} x' \frac{1+k_1 r^2 + k_2 r^4 + k_3 r^6}{1+k_4 r^2 + k_5 r^4 + k_6 r^6} + 2p_1 x' y' + p_2(r^2 + 2x'^2) + s_1 r^2 + s_2 r^4 \\ y' \frac{1+k_1 r^2 + k_2 r^4 + k_3 r^6}{1+k_4 r^2 + k_5 r^4 + k_6 r^6} + p_1(r^2 + 2y'^2) + 2p_2 x' y' + s_3 r^2 + s_4 r^4 \end{bmatrix}$$

- Where k are radial distortion parameters, p are the tangential distortion parameters, and s are the thin prism distortion parameters.





Camera Parameters Demonstration



python
X = 0.00
Y = 0.00
Z = 0.00
alpha = 0.00
beta = 0.00
gamma = 0.00
fx = 300.00
fy = 300.00
s = 0.00
cx = 250.00
cy = 100.00
k1 = 0.00
k2 = 0.00
k3 = 0.00
k4 = 0.00
k5 = 0.00
k6 = 0.00
p1 = 0.00
p2 = 0.00
s1 = 0.00
s2 = 0.00
s3 = 0.00
s4 = 0.00

Reconstructing 3D Positions from 2D Images

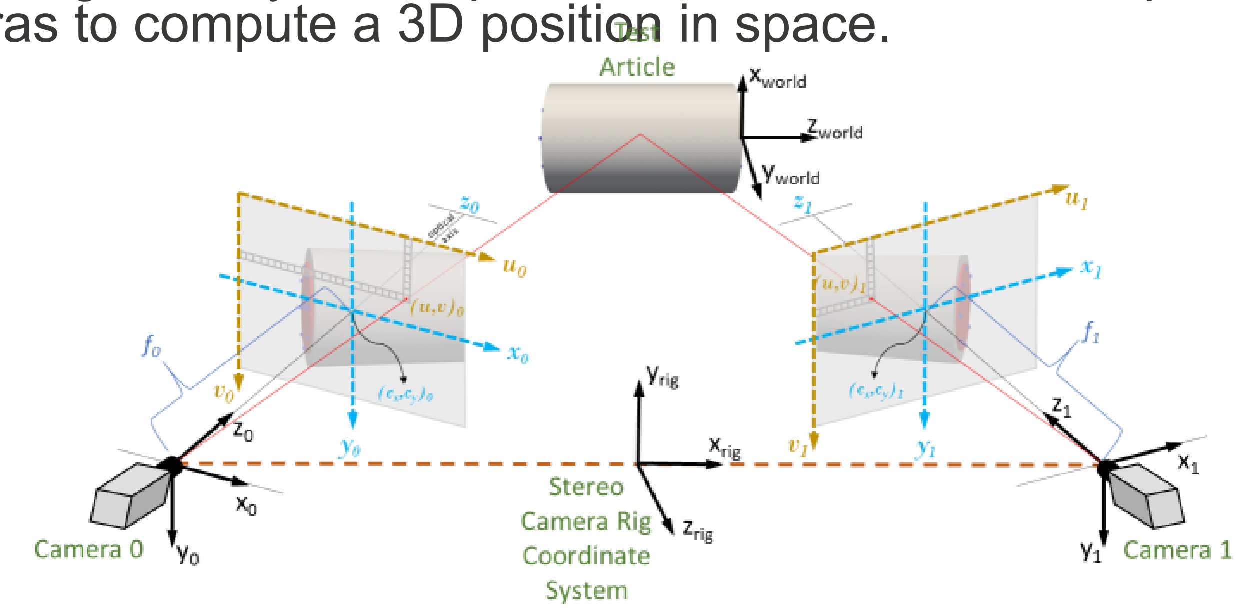
Triangulating Objects with Multiple Cameras

The camera equations result in a reduction of space, transforming a 3D point to a 2D point.

Obviously, some information is lost here, so we cannot simply reconstruct a 3D point from a position on a 2D image.

In general, a pixel on a camera image represents a line or ray in space originating from the camera pinhole and extending out through the camera imaging plane through the specified pixel.

We can then use geometry to compute the closest intersection point of rays from multiple cameras to compute a 3D position in space.



Epipolar Geometry

Epipolar geometry describes the geometry of a stereo imaging setup.

An **epipolar point** or **epipole** is the position of one camera's pinhole on the other camera's image.

All the points that correspond to a given pixel on the left image form a line across the right image. This is called an **epipolar line**.

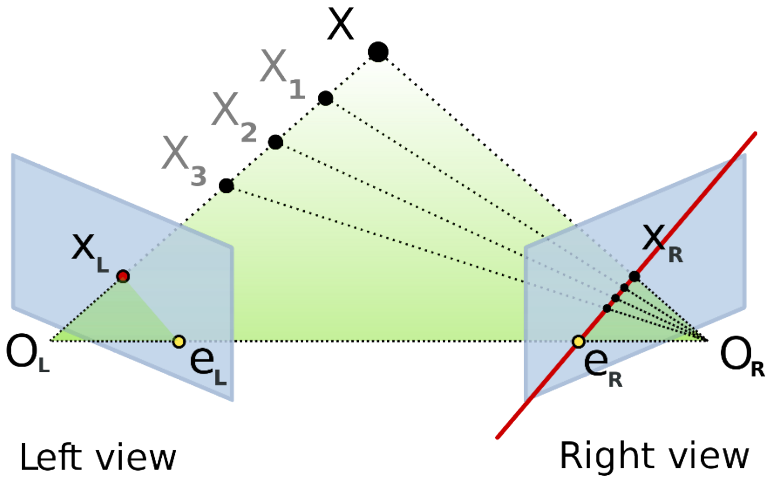
All epipolar lines pass through the epipole, but it is often off-image

- One camera doesn't typically image the other in a stereo setup

Epipolar lines depend on the position of a point in 3D space used to create that line.

An **epipolar plane** consists of the plane between the point in space and the two camera pinholes

- This plane's intersection with the image plane forms the epipolar lines.



Computing the line corresponding to a position on the image

A line in space can be defined by two points on that line.

- The first point will be the camera pinhole, trivially computed by $-\mathbf{R}^T\mathbf{T}$
- The second point must be found on the camera ray.

Imagining we are in the camera coordinate system, the intrinsic matrix determines the relationship between points in 3D space and points on the image.

$$c\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}f_x & s & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}X'\\Y'\\Z'\end{bmatrix}$$

Expanding the equations and normalizing the constant c, we get two equations for three unknowns, meaning a single point in space cannot be recovered.

$$u=\frac{f_xX'}{Z'}+\frac{sY'}{Z'}+c_x$$

$$v=\frac{f_yY'}{Z'}+c_y$$

However, we just need *any* position on the line, so we can assume e.g. $Z'=1$

$$\begin{bmatrix}u\\v\end{bmatrix}=\begin{bmatrix}f_x & s\\0 & f_y\end{bmatrix}\begin{bmatrix}X'\\Y'\end{bmatrix}+\begin{bmatrix}c_x\\c_y\end{bmatrix}$$

A position on the ray can be solved for by

$$\begin{bmatrix}X'\\Y'\end{bmatrix}=\begin{bmatrix}f_x & s\\0 & f_y\end{bmatrix}^{-1}\begin{bmatrix}u-c_x\\v-c_y\end{bmatrix}$$

$$Z'=1$$

We then transform to world coordinates using the inverse of the extrinsic camera matrix

$$\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}=\begin{bmatrix}\mathbf{R} & \mathbf{T}\\\mathbf{0} & \mathbf{1}\end{bmatrix}^{-1}\begin{bmatrix}X'\\Y'\\Z'\\1\end{bmatrix}$$

Closest point between two lines in 3D space

If two cameras are used, and a feature is found in both cameras' images, we can find the ray corresponding to that position in each image and then find the closest point between those rays

- If there were no errors, these rays should theoretically intersect, but small errors present in all experiments generally mean these rays will only come close to intersecting.

To find the closest point between two lines, we wish to find the line segment that is perpendicular to both lines.

We can express all points on a line in space $\mathbf{L}$ as

$$\mathbf{L} = \mathbf{P} + t\mathbf{N}$$

where $\mathbf{P}$ is a point on the line and $\mathbf{N}$ is the direction of the line.

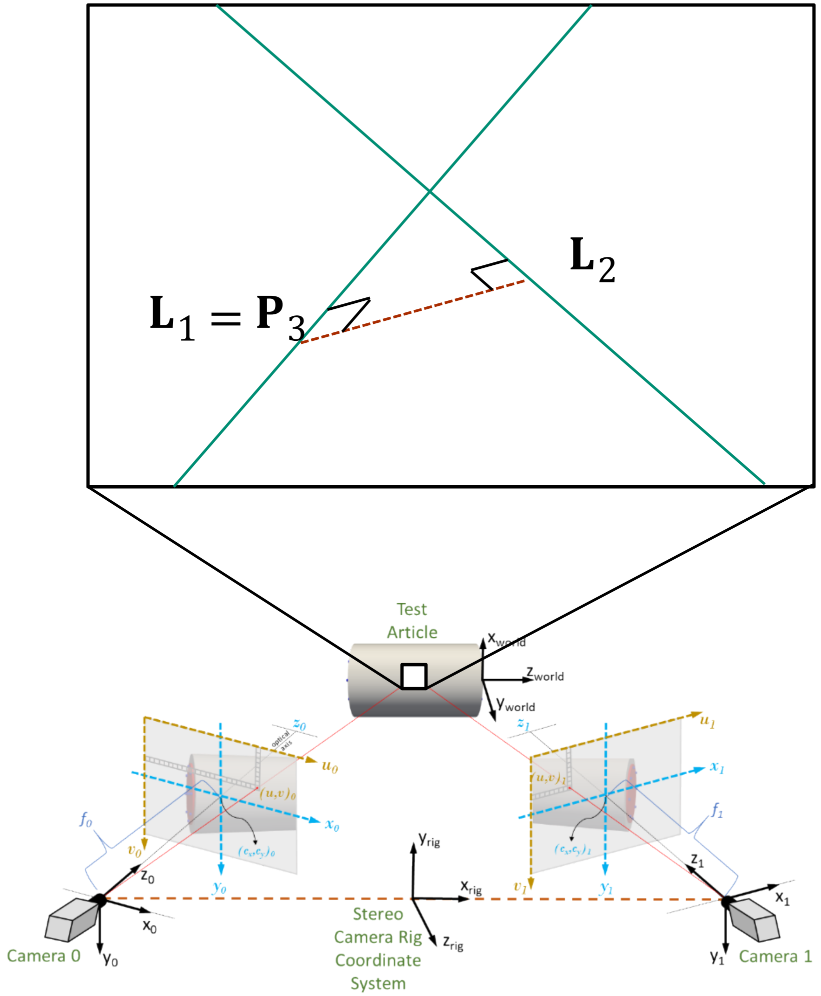
Given two lines, $\mathbf{L}_1 = \mathbf{P}_1 + t_1\mathbf{N}_1$ and $\mathbf{L}_2 = \mathbf{P}_2 + t_2\mathbf{N}_2$, we seek to find a third line $\mathbf{L}_3 = \mathbf{P}_3 + t_3\mathbf{N}_3$ that is perpendicular to the first two.

$$\mathbf{N}_3 = \mathbf{N}_2 \times \mathbf{N}_1$$

We then need to have the third line meet the first two lines. We assume that the third line starts from point $\mathbf{P}_3 = \mathbf{P}_1 + t_1\mathbf{N}_1$ on the first line and meets line two at $\mathbf{P}_2 + t_2\mathbf{N}_2$ after travelling distance t_3 in the direction $\mathbf{N}_3$.

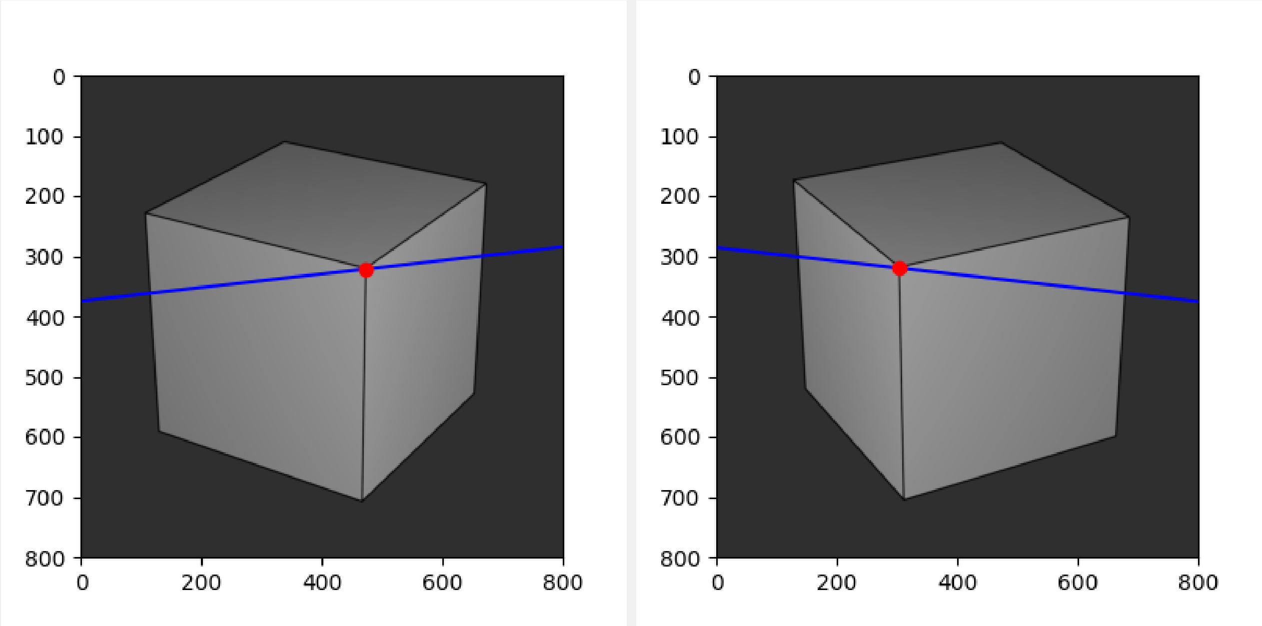
$$\mathbf{P}_1 + t_1\mathbf{N}_1 + t_3\mathbf{N}_3 = \mathbf{P}_2 + t_2\mathbf{N}_2$$

Which forms three equations for three unknowns t_1, t_2, t_3 and allows us to solve for the closest point on line 1 to line 2 and vice versa. These two points can be averaged to get the mean point between the two lines.



Triangulation Demo

python




```
Left Camera:
K:
[[2222.222    0.      400.   ]
 [   0.     2222.222  400.   ]
 [   0.        0.        1.   ]]
RT:
[[ 0.851  0.525  0.    -0.008]
 [ 0.234 -0.379 -0.895 -0.06 ]
 [-0.47   0.762 -0.445 11.256]]
Selected Pixel:
[472.727 322.078]

Intersection Point:
[ 0.994 -0.999  0.992]
Intersection Accuracy:
0.0016878214701343 84
```

```
Left Camera:
K:
[[2222.222    0.      400.   ]
 [   0.     2222.222  400.   ]
 [   0.        0.        1.   ]]
RT:
[[ 0.474  0.88  -0.    -0.008]
 [ 0.392 -0.211 -0.895 -0.06 ]
 [-0.788  0.425 -0.445 11.256]]
Selected Pixel:
[303.896 319.481]
```

Camera Calibration

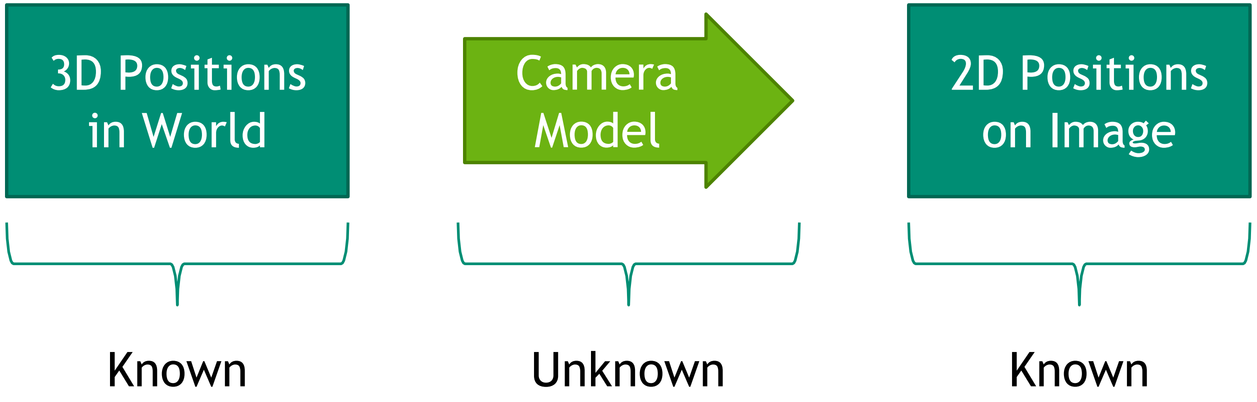
Constructing Camera Parameters from Images

Manufacturer quantities for focal length and other lens parameters are generally rough approximations.

It is also difficult to exactly measure a cameras position and orientation.

Therefore we cannot generally construct accurate camera equations without performing some kind of camera calibration.

Camera calibration attempts to fit a pinhole camera model and lens distortion model to a set of images containing features with known geometry.



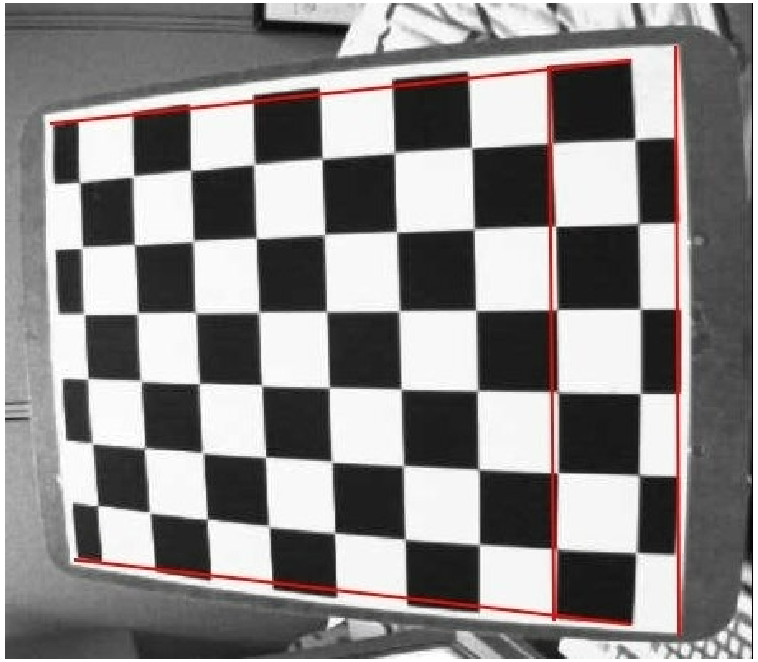
Images with Known Geometry

For calibration images, we generally acquire images of an object with known geometry, for example:

- Checkerboard
- Dot-Grid

Camera calibration software is able to extract features from calibration images automatically to populate the positions of the features on the image

The user generally provides some kind of geometry information to the calibration software (grid spacing, etc.) to let the calibration software populate the positions of the features in

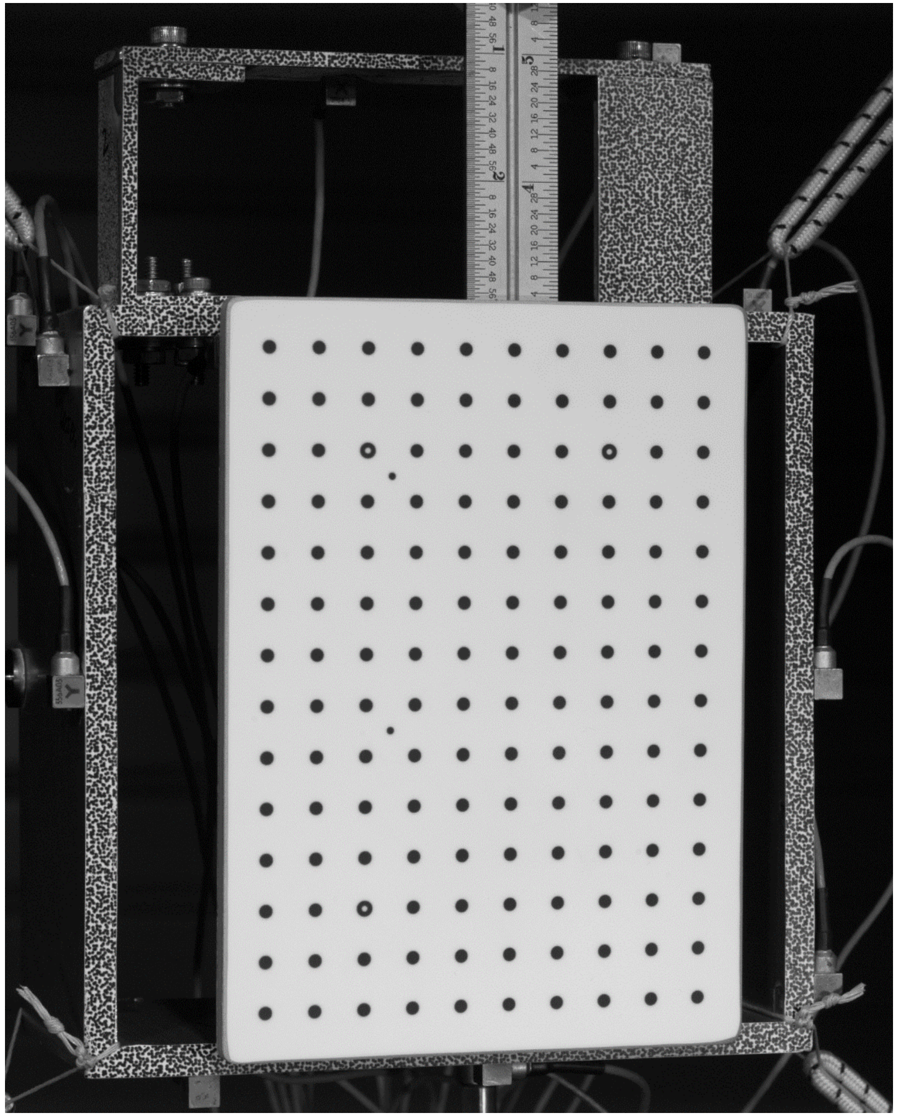


Calibration Volume

Most camera calibration routines accept a large number of images of the calibration target in multiple orientations and positions.

The volume of space that the features on the calibration target traverse is called the calibration volume.

Generally, the calibration volume should contain the volume that the test article should occupy, otherwise you end up extrapolating the calibration rather than interpolating.



Not best practice!

Solving for the Camera Parameters

Bundle adjustment is the problem of simultaneously updating 3D geometry, relative motion between images, and parameters of the camera model given a set of images of the scene.

Camera calibration with a well-defined calibration target is a more-constrained, better-behaved version of bundle adjustment, because the 3D geometry of the scene is generally known.

Effectively an optimization problem on the camera intrinsic and extrinsic properties to minimize reprojection error of the features on the image.

Camera equations are nonlinear, so this is generally an iterative process using a nonlinear least-squ … n.

$$\min \Sigma\left(\text{Predict 2D Positions on Image} - \text{Actual 2D Positions on Image}\right)$$

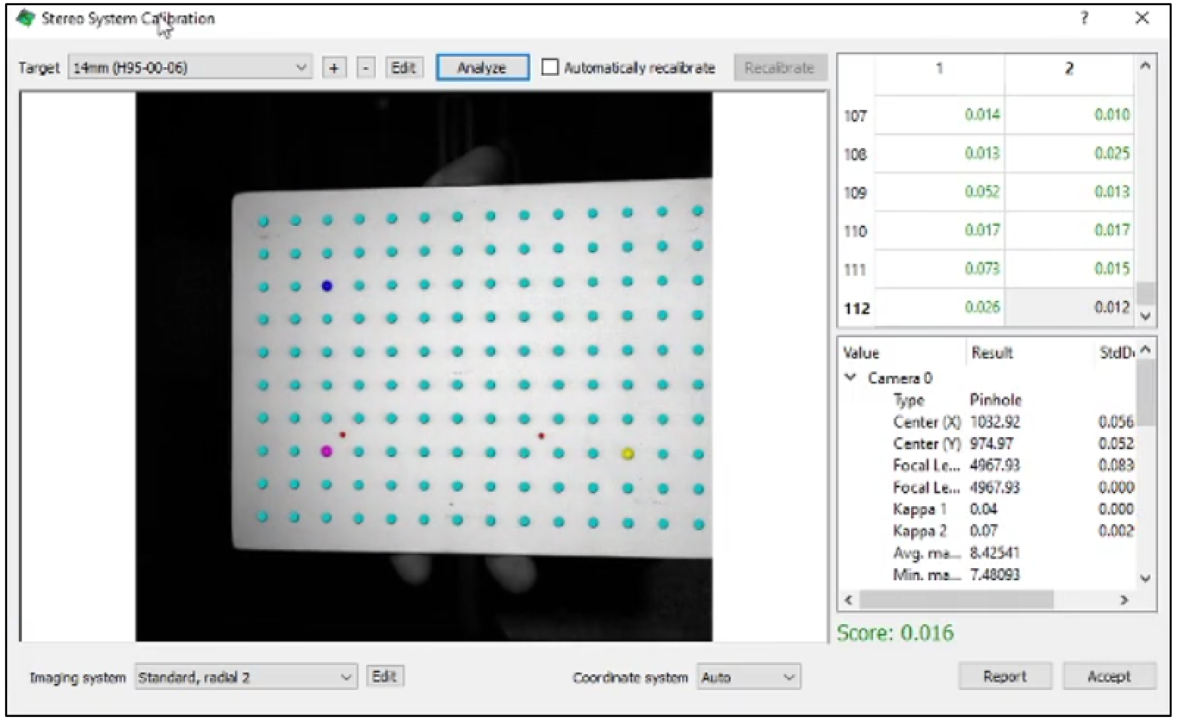
Evaluating Camera Calibration Results

Camera calibration routines will report a reprojection error metric, which is a measure of how far away features on the image are from where they are predicted to be by the camera model.

- Often times taken as an RMS over all features in all calibration images.
- Different software packages might have different metrics, so it is important to understand what values are acceptable

Calibration software will also generally have a way to remove bad points or images

- Images where features are on the edge of the image
- Images where there is blur or depth of field issues
- Images with glare or other lighting artifacts that make feature extraction inaccurate



Calibration Dialog from Correlated Solutions’ VIC3D Software

Selecting a Lens and Camera Combination for a Given Test

Selecting Cameras and Lenses for Optical Testing

I have a test article on which I'd like to perform some optical measurements, which cameras and lenses should I use?

Optimal camera and lens selection will depend on a variety of constraints in the imaging setup.

Things to think about:

- Geometry:
  - Stand-off distance – Are there constraints on where I can position my cameras?
  - Camera orientation – What surface(s) am I trying to image? Consider depth-of-field requirements when viewing at oblique angles
  - Test article size – How does the size of the test article constrain camera positions given the lenses I have available?
- Camera:
  - Image Resolution – What level of detail is required? How many pixels of displacement are expected? How much data will I generate?
  - Sensor size – Modifications on lens view angle/zoom
  - Noise Level – Will I be shooting in low-light or otherwise suboptimal exposure conditions?
  - Sample rate – What is the bandwidth of the test? How many images per second are required? Does that constrain image resolution?
- Lens:
  - Nominal focal length – What is the view angle of my lens, can I see the whole test given camera positioning constraints?
  - Minimum focal distance – What is the closest the camera can be to the test article?
  - Aperture – What is the maximum aperture of the lens? What is the required depth of field? Will there be issues with diffraction?

Note that there can be multiple setups that can yield satisfactory results.

- We often use an iterative process to figure out all configurations that work and then select our favorite.

Computing the Imaging Volume

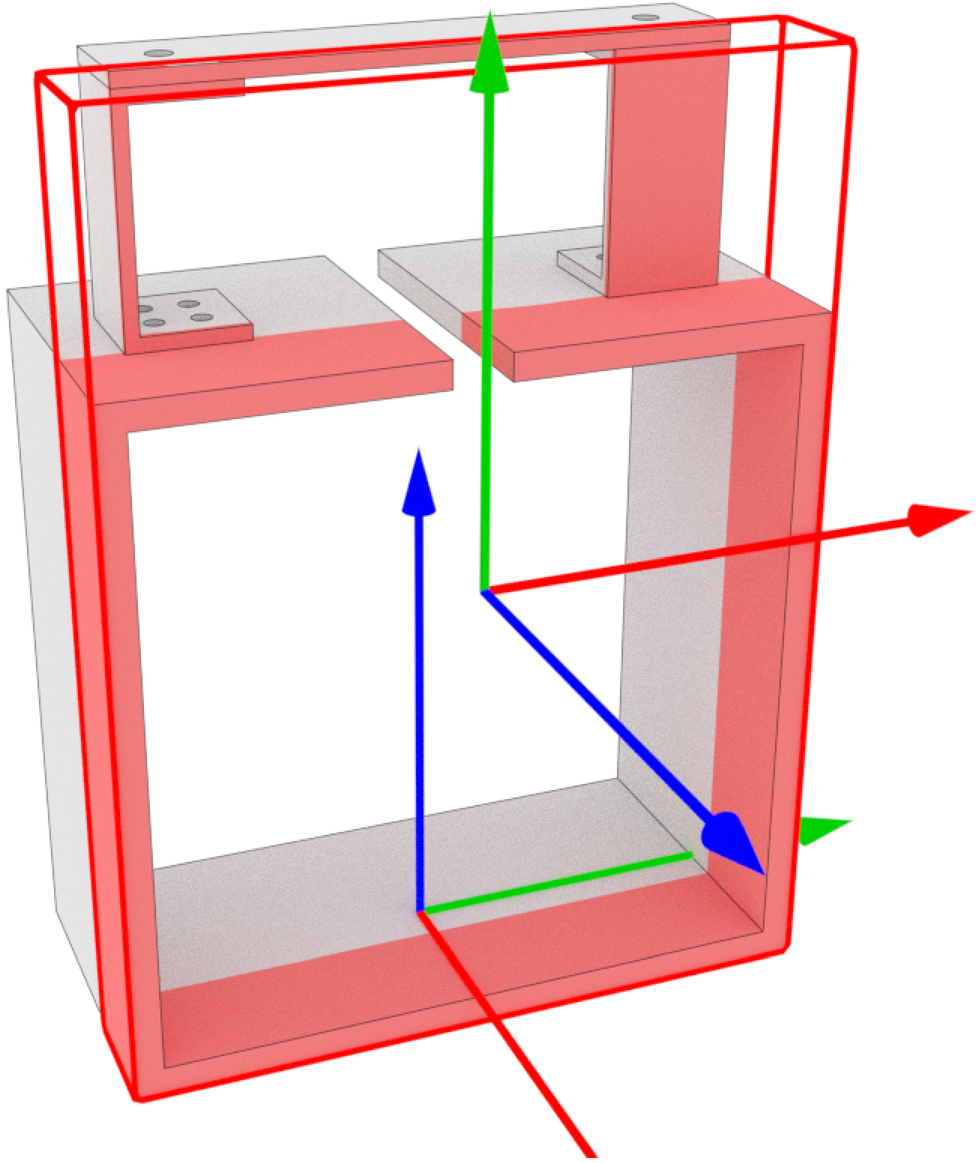
To ensure our cameras can image our test, we first need to figure out what portion of the test article we want to image.

For rough calculations, we set up a bounding box around the portions of the test article we will image.

Note that this does not need to contain the entire test article, only the portions of the test article that will be imaged.

Note that we can transform the bounding box coordinates to image coordinates using the camera projection equations to determine if the bounding box fits entirely on the camera image for a given setup.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{R}|\mathbf{T}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



Selecting a Camera and Lens

The intrinsic parameters of the camera must be determined prior to being able to place the camera in 3D space.

$$\mathbf{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

As a first approximation:

- c_x and c_y can be set to half the image resolution in the vertical and horizontal directions, respectively.
- f_x can be set equal to f_y, which is equal to the lens focal length in mm divided by the pixel size in mm.
- s can be set to zero.

For our example, we've chosen an 85 mm lens with a camera sensor that is 2048 x 1920 pixels over 27.6 x 25.9 mm, resulting in 74.2 pixels/mm. For an uncropped image with this setup, the intrinsic matrix **K** will be

$$\mathbf{K} = \begin{bmatrix} 6307.2 & 0 & 1024 \\ 0 & 6307.2 & 960 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that the real camera/lens system will have slightly different values than these due to:

- Focus breathing – generally focal lengths are defined when focus is at infinity
- Lens distortions – barrel and pincushion distortions will modify the edges of the image slightly
- Rounding in manufacturer's reported focal length – often rounded to the nearest 5 or 10 mm

Upon performing a camera calibration, the true focus value $f = 6890.5$, meaning the lens was actually a 92 mm focal length.

This just serves as a rough estimate so we can look at the feasibility of using a specific lens.

Lens: Zeiss Milvus 85mm f/1.4

Camera: Phantom v2640

Positioning and Orienting the Camera

Some considerations for camera placement:

- We want the camera to be able to see the surfaces of interest
- We want the camera to put the most pixels as possible across our test article (i.e. zoom in as much as possible with margin for test article motion)
- Orientations might be constrained due to a stereo camera setup

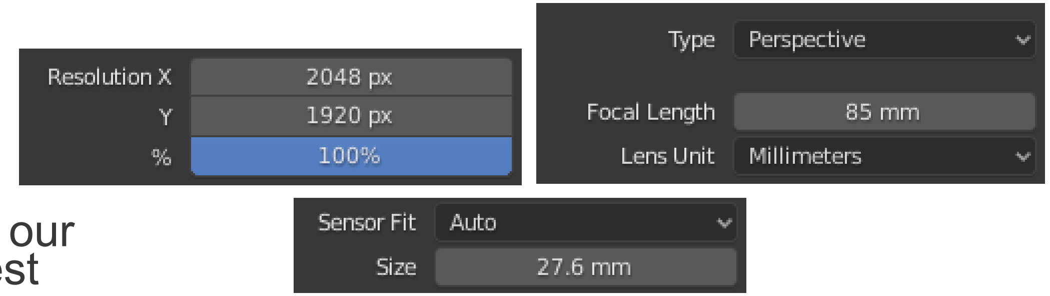
Positioning the camera can be an iterative process.

- Place camera in space
- Project bounding box to image via camera equations
- Check if all corners of bounding box are within the image
- Repeat as necessary

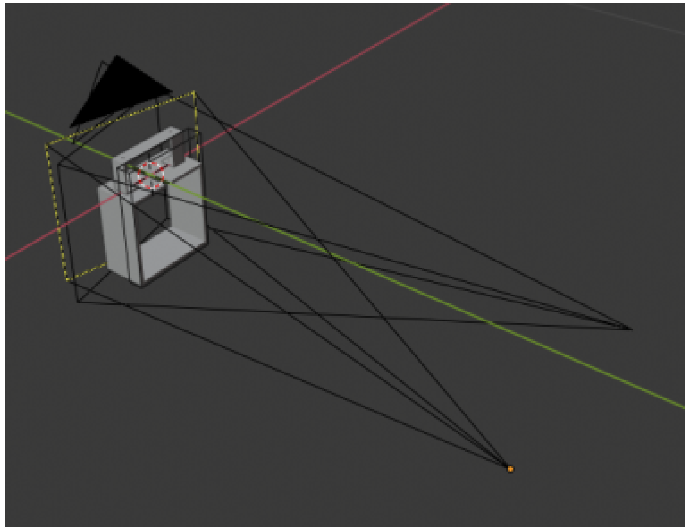
Cameras can also be placed interactively using rendering software such as Blender

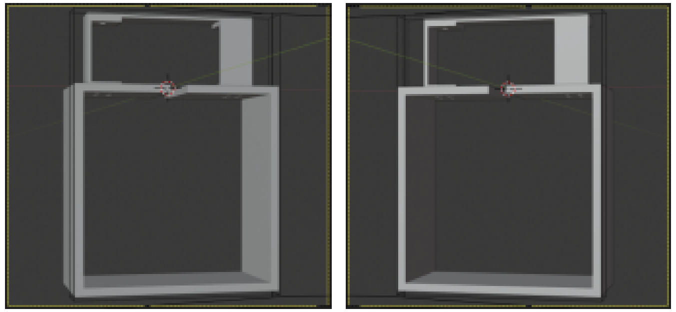
- Set up camera’s focal length, sensor size, and image resolution
- Position camera so that it fills the image.
- Extract the camera matrices from parameters in Blender

Ensure that the lens is not placed closer to the part than its minimum focus distance.



Blender Camera Parameters





Cameras positioned in Blender and resulting perspective

Constructing [R|T] from camera position and orientation

Often times, cameras are specified with a location $\mathbf{T}_{cam}$ and orientation $\mathbf{R}_{cam}$ in 3D space

- This is the transformation that takes the camera coordinate system and transforms it to the world coordinate system
- Transforms the camera pinhole from (0,0,0) in the camera coordinate system to $\mathbf{T}_{cam}$ in the world coordinate system
- **[R|T]** in the camera equations is the inverse transformation, transforming world coordinate system to the camera coordinate system

$$\mathbf{R} = \mathbf{R}_{cam}^{T}$$
$$\mathbf{T} = -\mathbf{R}_{cam}^{T}\mathbf{T}_{cam}$$

or

$$\begin{bmatrix} & \mathbf{R} & & \mathbf{T} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} & \mathbf{R}_{cam} & & \mathbf{T}_{cam} \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1}$$

*Note for Blender users: Blender positions its cameras with its optical axis along the Z- axis of the camera coordinate system rather than the Z+ axis, so an additional 180° rotation about the camera X axis is required.

Imaging Checks – Depth of Field

We can do rough calculations to determine the depth of field we will achieve with a given lens and aperture.

$$DOF_{near} = \frac{Hd}{H + d}$$

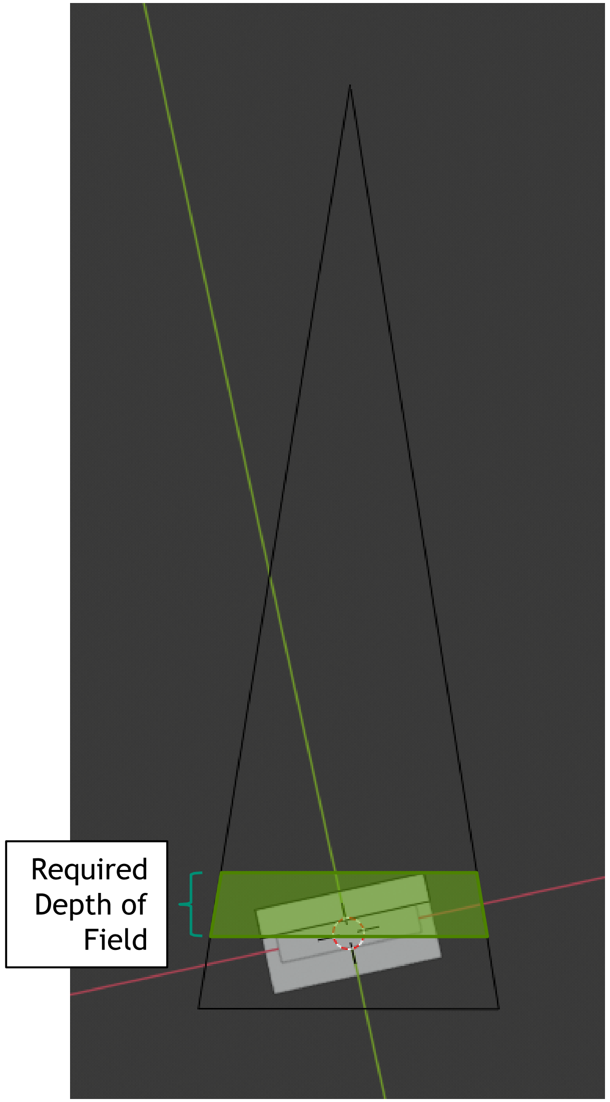
$$DOF_{far} = \frac{Hd}{H - d}$$

Here d is the focus distance of the lens, and H is the "hyperfocal distance" of the lens, defined as the focus distance for a lens where everything past this distance is in focus

$$H = \frac{f_{mm}^2}{Nc}$$

Where f_{mm} is the focal length of the lens in millimeters, N is the f-number (e.g. 8 if the aperture is $f/8$) and c is the "circle of confusion" which can be conservatively set to the size of one pixel on the sensor

When focusing at hyperfocal distance, $H = d$ so $DOF_{near} = \frac{d^2}{2d} = \frac{1}{2}d$ and $DOF_{far} = \frac{d^2}{0} \to \infty$



Solving for the required aperture for a given lens

To determine whether the depth of field is adequate, we need to determine how far each point is from the camera's focus plane.

To do this, we transform the bounding box coordinates into the camera coordinate system

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} & T_x \\ R_{yx} & R_{yy} & R_{yz} & T_y \\ R_{zx} & R_{zy} & R_{zz} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

We then pick out the maximum and minimum Z' values and set the focal distance d to be halfway between the two.

We can then solve for the aperture values that give in-focus values at the closest and farthest values of Z'

$$N_{near} = \frac{f^2(d - DOF_{near})}{c\,d\,DOF_{near}}$$

$$N_{far} = \frac{f^2(DOF_{far} - d)}{c\,d\,DOF_{far}}$$

Imaging Checks – Diffraction

A smaller aperture results in a larger depth of field, but it can also result in diffraction effects.

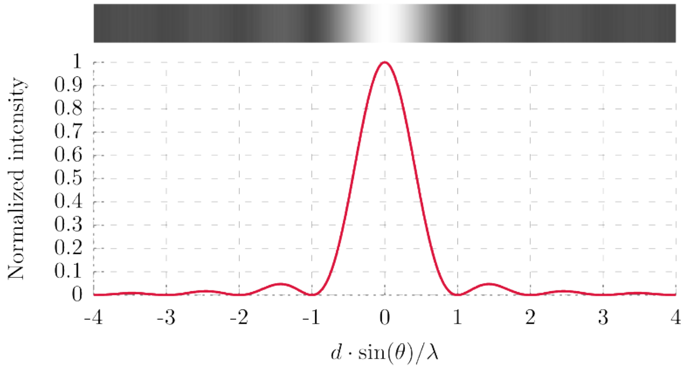
For an optical system with a circular aperture, a point is not imaged as a point, but rather an Airy disk, with one central lobe and sequentially degrading side-lobes.

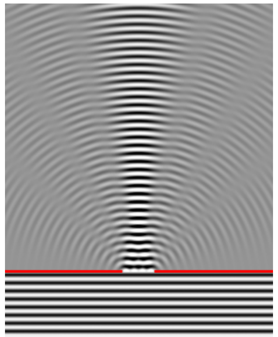
The diameter of this disk is approximately

$$D = 2.44\lambda N$$

Where λ is the wavelength of light and N is the f-number of the aperture of the lens.

An optical system is called "diffraction limited" when the diameter of the Airy disk is larger than the size of a pixel on the sensor; no more sharpness can be obtained from the lens!





Review

Camera Concepts

The lens and the sensor combine to determine the field of view of an image.

- Longer focal length and smaller sensors result in a more “zoomed in” image

The brightness of an image is a function of:

- Exposure time: can result in motion blur if too long
- Aperture Size: larger aperture is brighter but limits depth of field
- Image Gain/ISO: keep at base levels unless absolutely necessary
- Lighting in the scene: keep in mind heat generated by lights

Generally better to slightly underexpose an image than overexpose an image

- Similar to data acquisition system, don’t want to clip the signals

Camera Geometry

The Pinhole Camera Model is ubiquitous in optical testing, representing the transformation between world coordinates and image coordinates.

$$c\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}f_x & s & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}R_{xx} & R_{xy} & R_{xz} & T_x\\R_{yx} & R_{yy} & R_{yz} & T_y\\R_{zx} & R_{zy} & R_{zz} & T_z\end{bmatrix}\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}=\mathbf{K}[\mathbf{R}|\mathbf{T}]\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}$$

The Pinhole Model is typically supplemented with lens distortion models, which have some assumed functional form

A position on an image represents a line through 3D space, so 2 cameras are needed to triangulate a location in space.

Triangulation can be treated as the closest intersection between two lines.

Camera parameters are fit using camera calibration procedure, which typically occurs using a dedicated calibration target with known geometry.

Selecting a Camera and Lens for a Given Test

Camera considerations:

- Sample rate vs. image resolution
  - Expected displacements drive resolution requirements
  - Structure natural frequencies drive sample rate requirements
  - Maximum shutter time limited by the sample rate of the camera
- Sensor size
  - Smaller sensors will be more zoomed in
  - Smaller sensors may have more noise due to smaller, more closely packed electronics and less light collected
- Position and orientation
  - Ensure the test article is visible in the image
  - Ensure not closer than the minimum focal distance of the lens

Lens considerations:

- Focal length
  - Larger values will be more zoomed in
- Aperture
  - Consider depth of field vs. exposure required
  - Check for diffraction limiting your image
- Minimum focus distance
  - Ensure your part can be put in focus