# End to End:  Part 1

Earl Wong

# Subjects

- **Optics**
- Sensor
- ISP
- GPU
- NPU

# Optics - Image Formation

- We begin with image formation.
- We explain the basics, using a pinhole camera.
- A pinhole camera uses the following constructs to create an image:

1) ray tracing

2) a small “pinhole” aperture

3) an underlying scene

# Optics - Image Formation

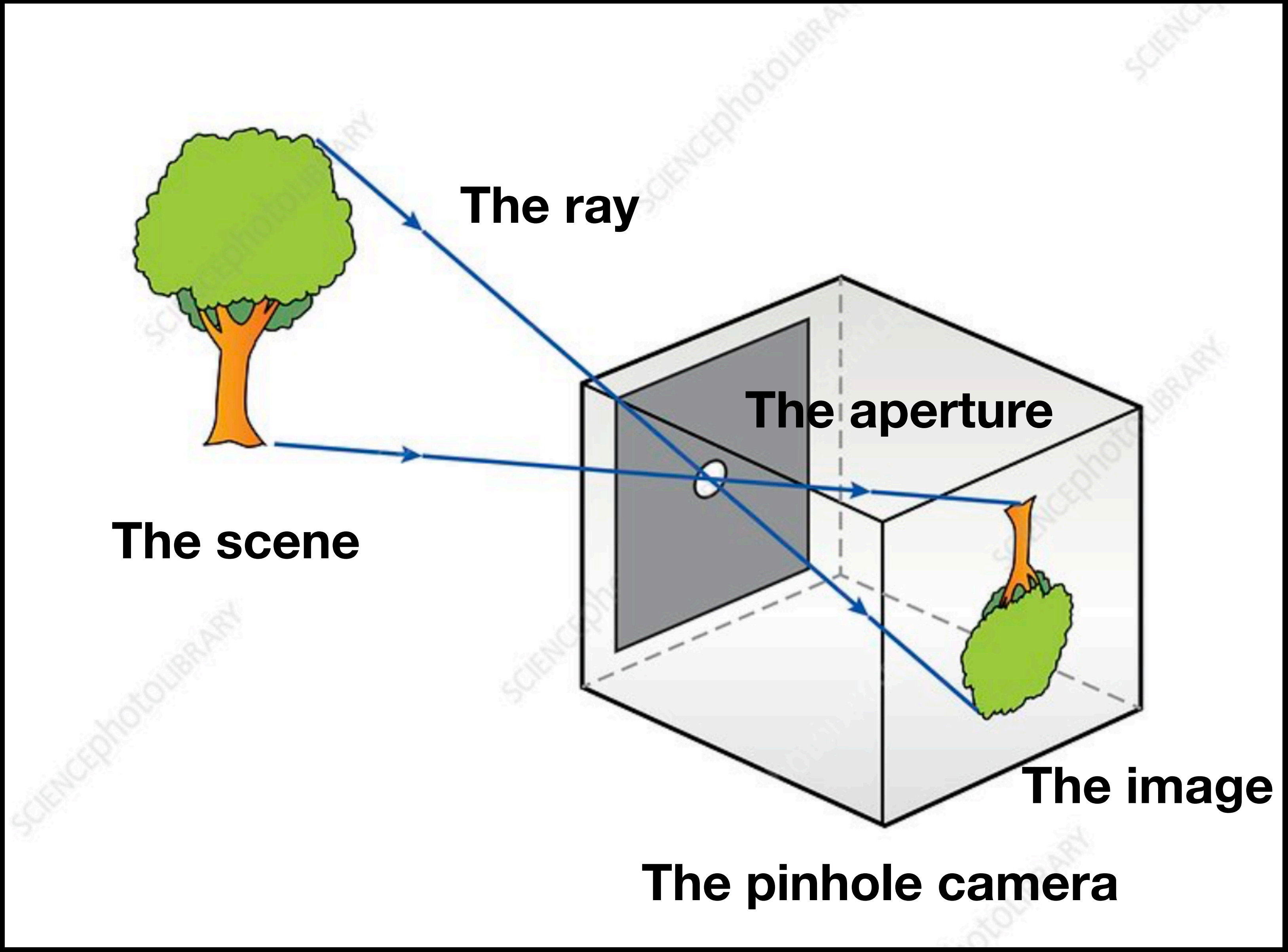
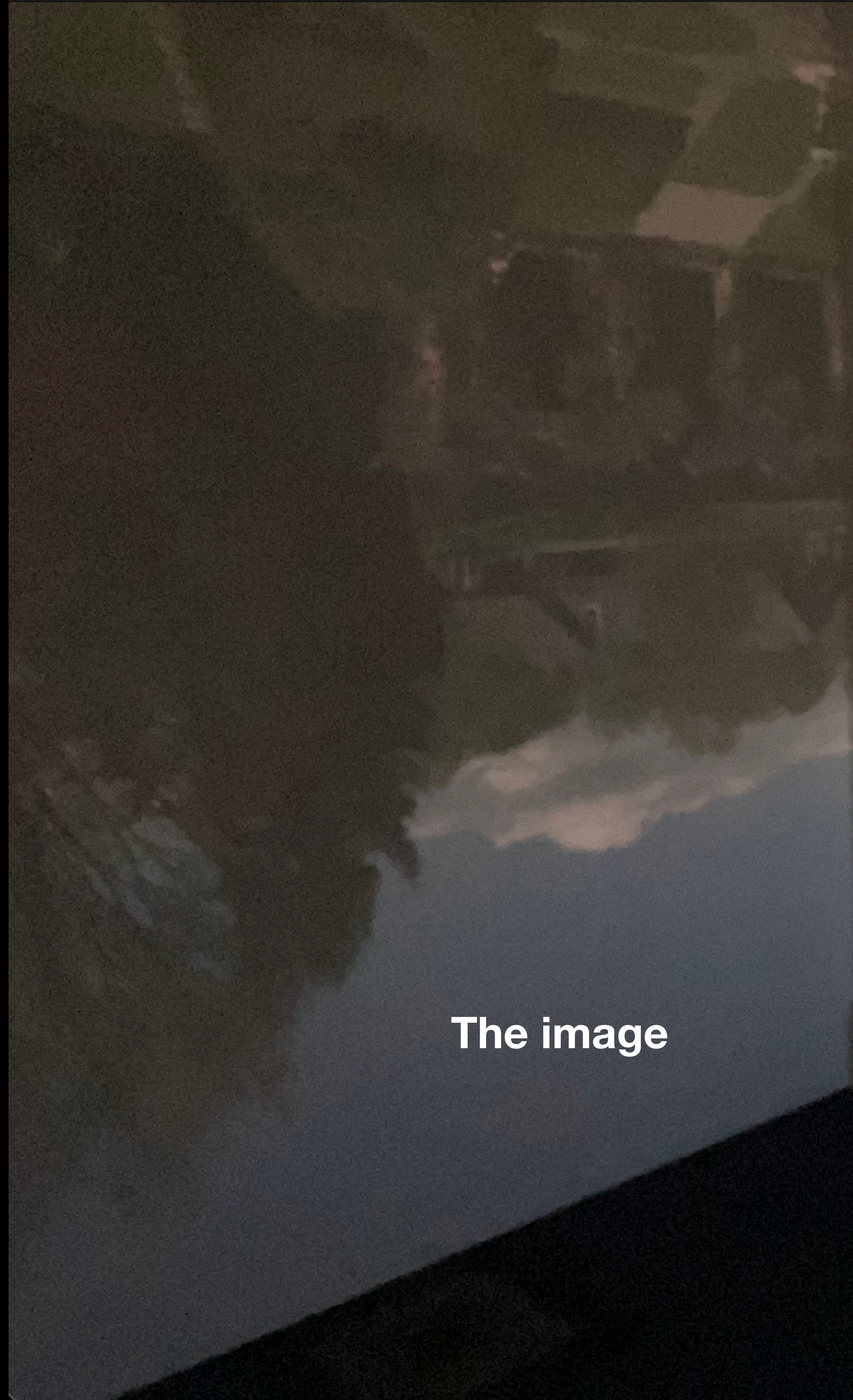



Image courtesy of Science Photo Library

# Optics - Image Formation

- Using ray tracing, we can map illumination sources from the scene into the image plane.
- The next two slides show a pinhole camera in operation, taken by the author at the Eastman Kodak Museum, in Rochester, New York.

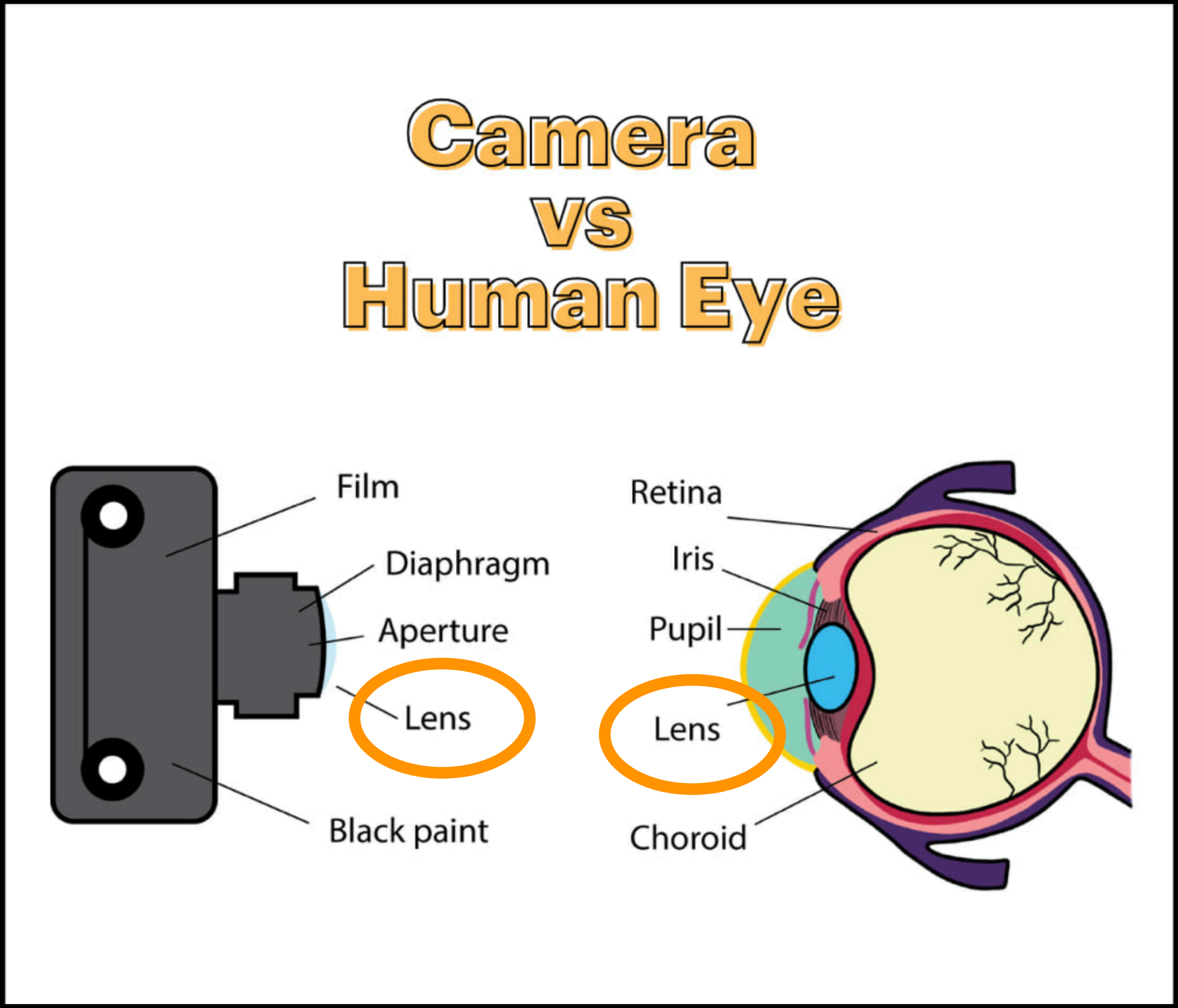

The aperture


The image

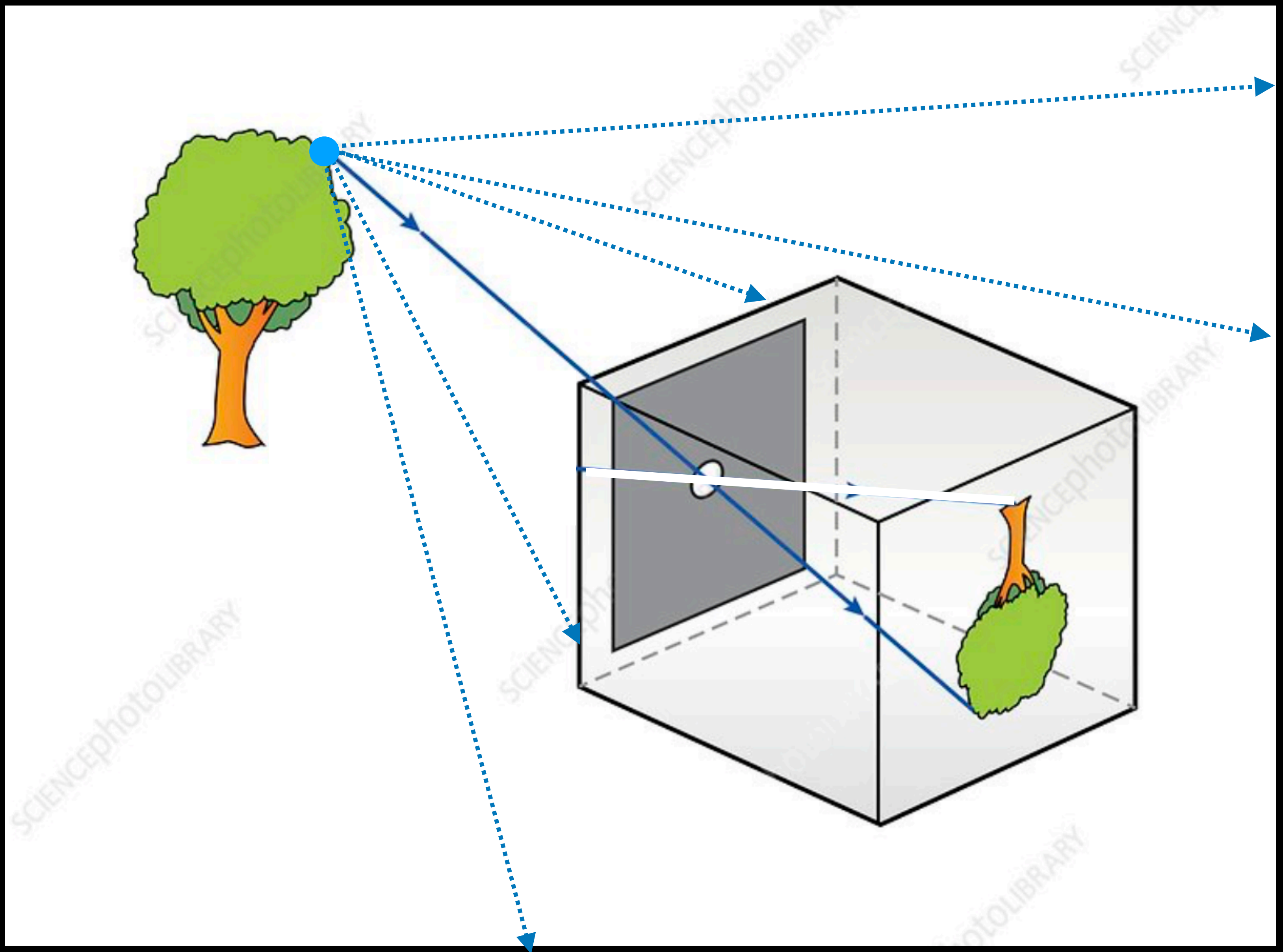
# Optics - Image Formation

- The human visual system is an advanced pinhole camera - a camera with additional “features”.
- The human visual system has a lens for gathering and focusing light.
- The human visual system has a retina (containing photoreceptors) for capturing the image formed by the light.
- In this presentation, we will focus on the lens, and show how lens design is closely tied to the image sensor specifications.
- In the next presentation, we will talk about image sensors in more detail.

# Optics - Image Formation



# Optics - Image Formation


SCIENCEphotoLIBRARY

**By construction, every point in the scene emits numerous rays.**
**However, a pinhole camera only allows a small number of these rays to be imaged.**

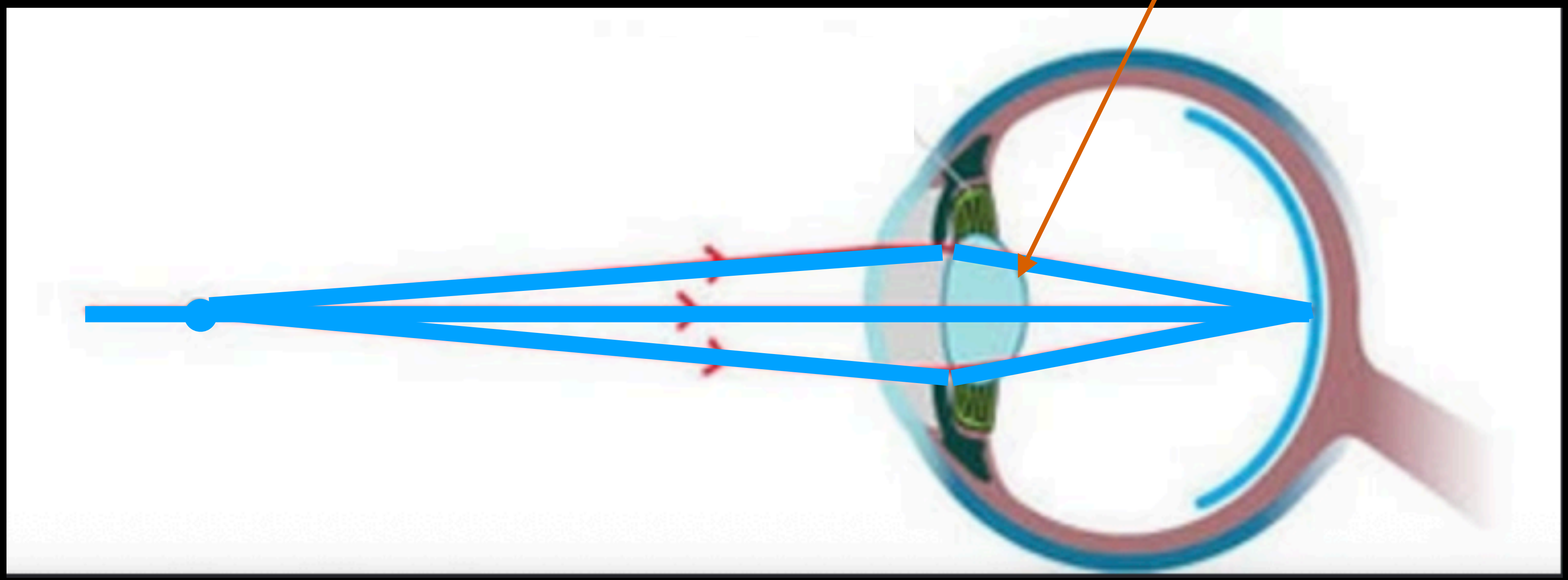
**Once a lens is substituted for the pinhole aperture, 1) additional rays can be collected and 2) focused.**

# Optics - Image Formation

- To increase the light gathering capabilities, a lens is introduced.
- The lens “collects” rays emitted from the scene location / point source, redirecting them to the image plane.
- In addition, the lens also focuses the rays onto the image plane.
- More will be said on the latter, during the discussion on depth of field (DoF).

# Optics - Image Formation

Lens of the human eye



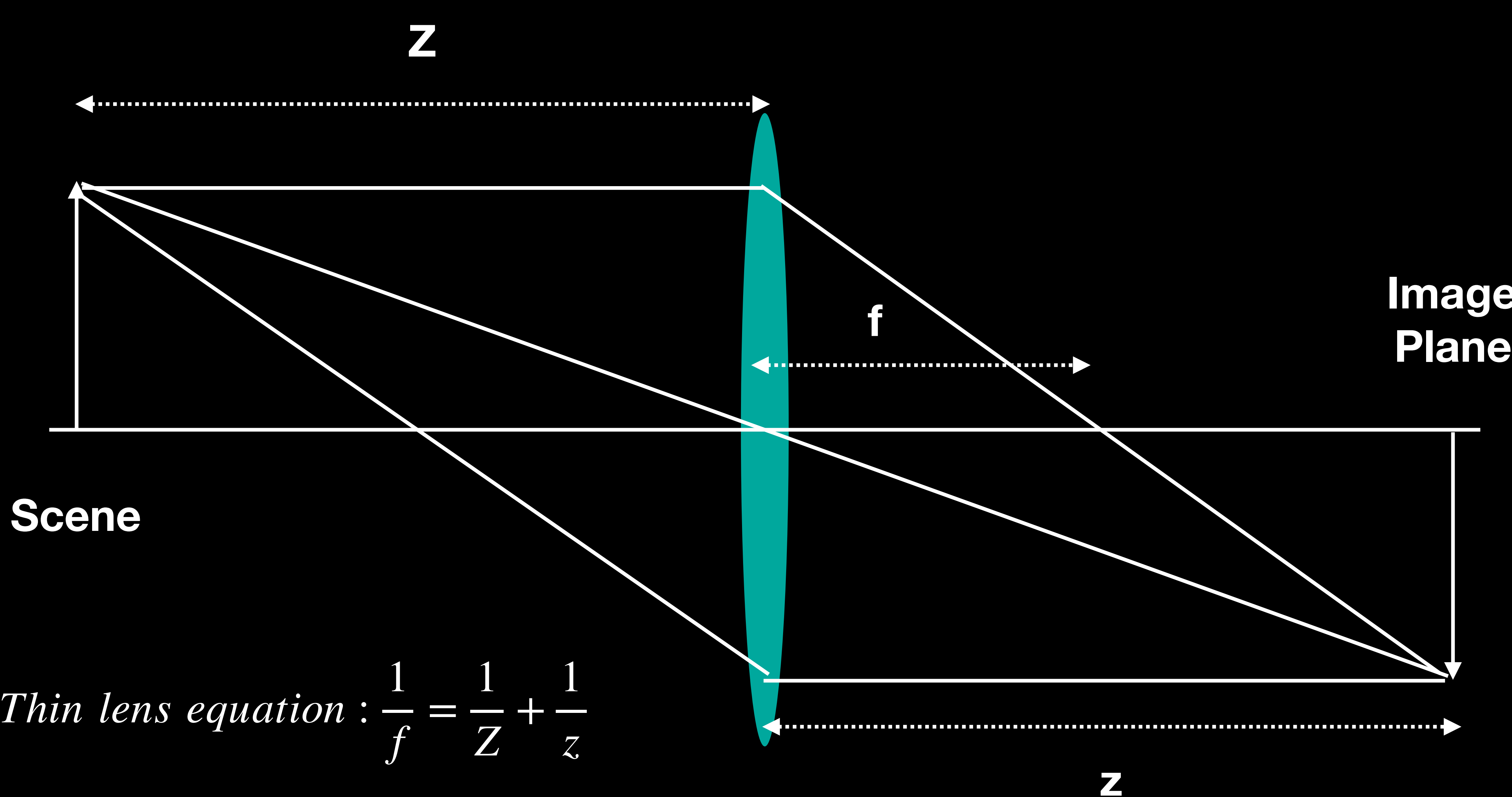
The human visual system collects multiple rays emitted by the source, focusing the rays onto the retina.

# Optics - Image Formation

- The simplest lens model is the thin lens model.
- The thin lens model is governed by the following

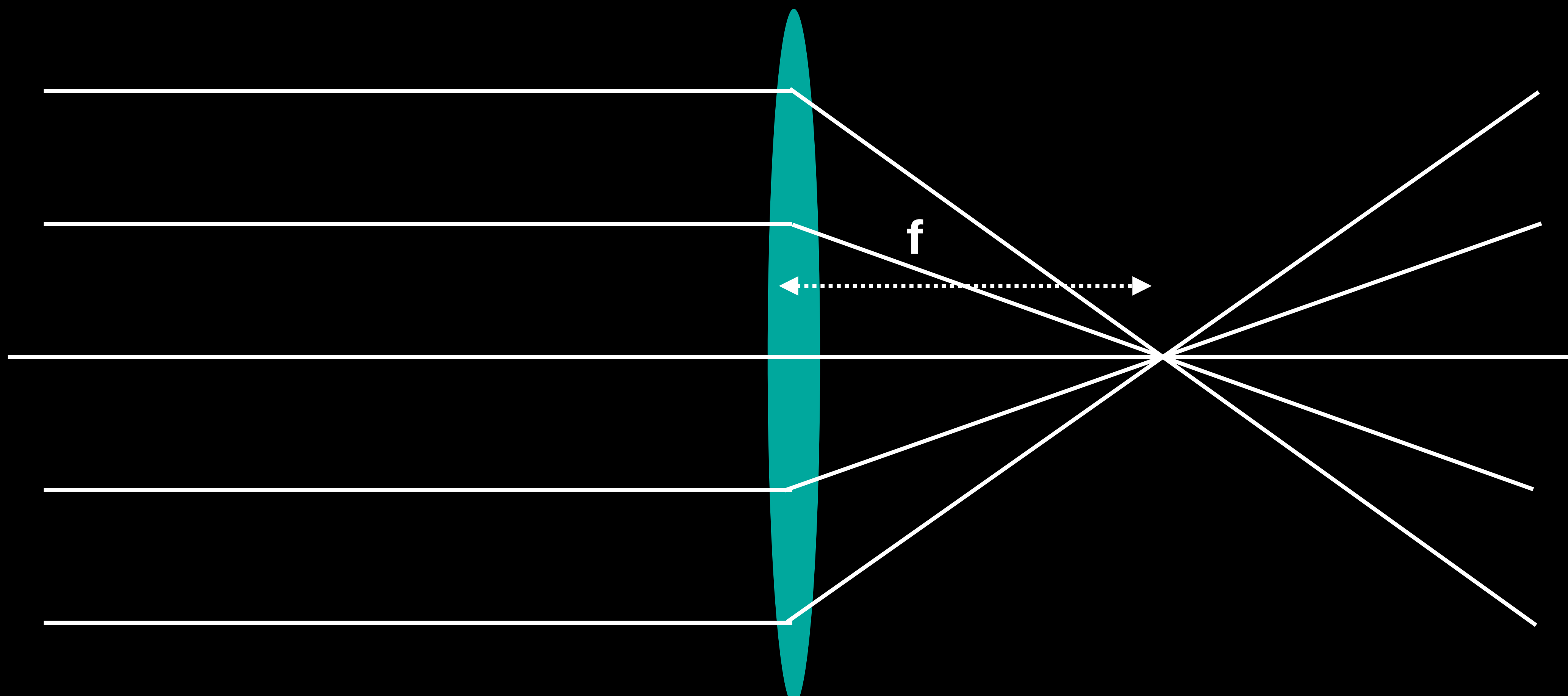
  equation: $\frac{1}{Z} + \frac{1}{z} = \frac{1}{f}$
- The thin lens equation says the following: For an object at a distance Z from the thin lens, the image will form at a distance z from the thin lens, for a thin lens with focal length f.
- The thin lens equation can be derived using ray tracing and triangle equalities.
- The lens focal length f is defined to be the location on the optical axis where parallel rays from the scene (=very distant object point source) and perpendicular to the lens, are focused at a single point.

# Optics - Image Formation




$$Thin\ lens\ equation : \frac{1}{f} = \frac{1}{Z} + \frac{1}{z}$$

# Optics - Image Formation




**focal length f of a lens**

# Optics - Image Formation

- But what happens, if various objects in the scene are at different distances, like in real life?

- Objects at one distance will be in focus, while objects at other distances may or may not be.

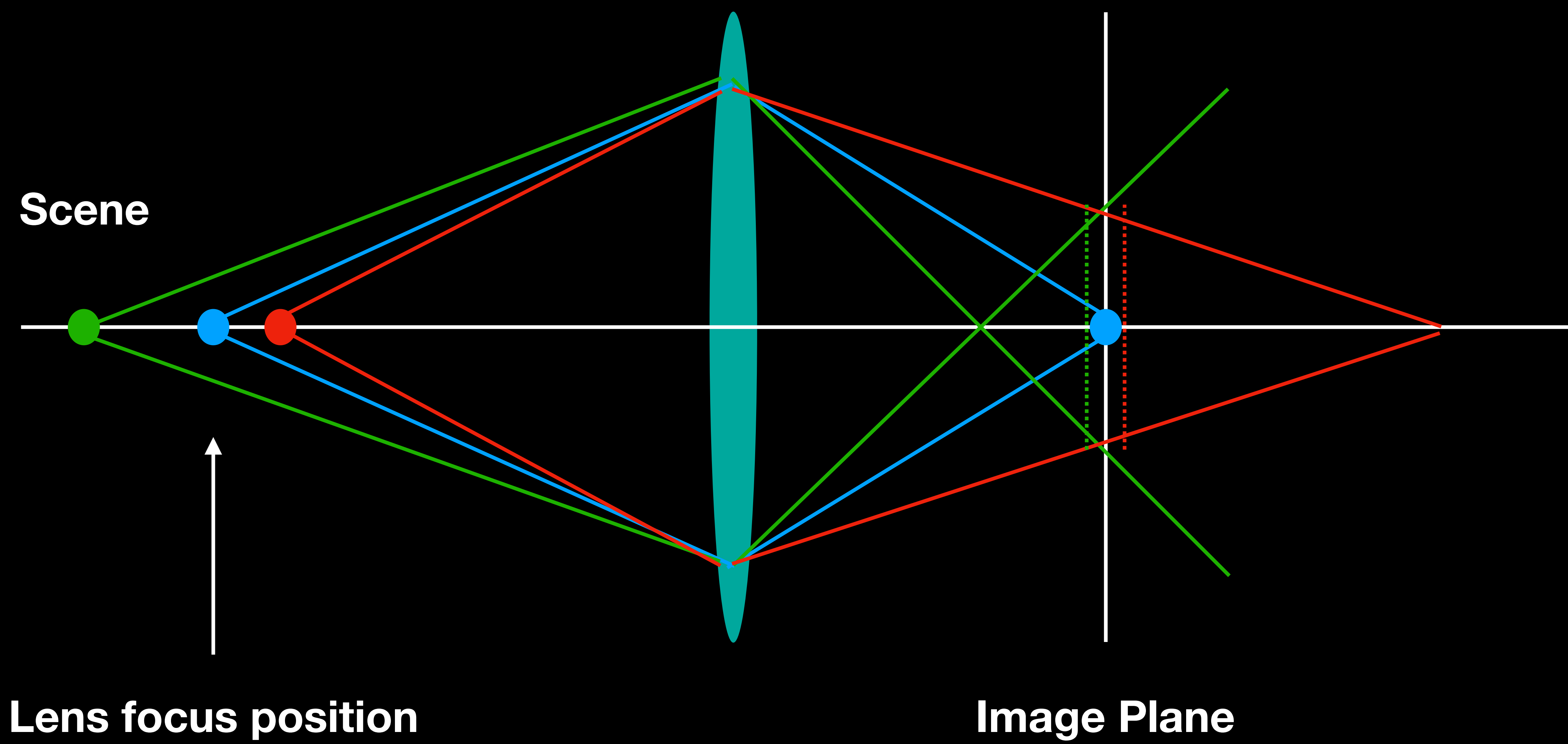
- To determine what is and is not in focus, the depth of field (DoF) needs to be computed.

- The DoF determines (for a given lens focus position), the “range” of distances in the scene that will be imaged “in focus” on the image plane.

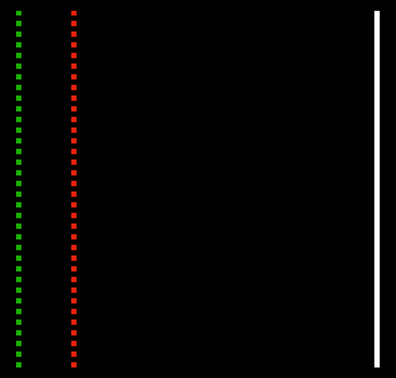
# Optics - Image Formation

- In practice, the DoF can vary, as a function of the focus position.
- i.e. The DoF is larger for distant lens focus positions and smaller for closer lens focus positions.
- The DoF is also a function of the lens focal length.
- i.e. The DoF is smaller for lenses with larger focal lengths and larger for lenses with smaller focal lengths.
- In addition, the DoF is also a function of the pixel size and the aperture size; we illustrate these two cases, using ray diagrams.

# Optics - Image Formation

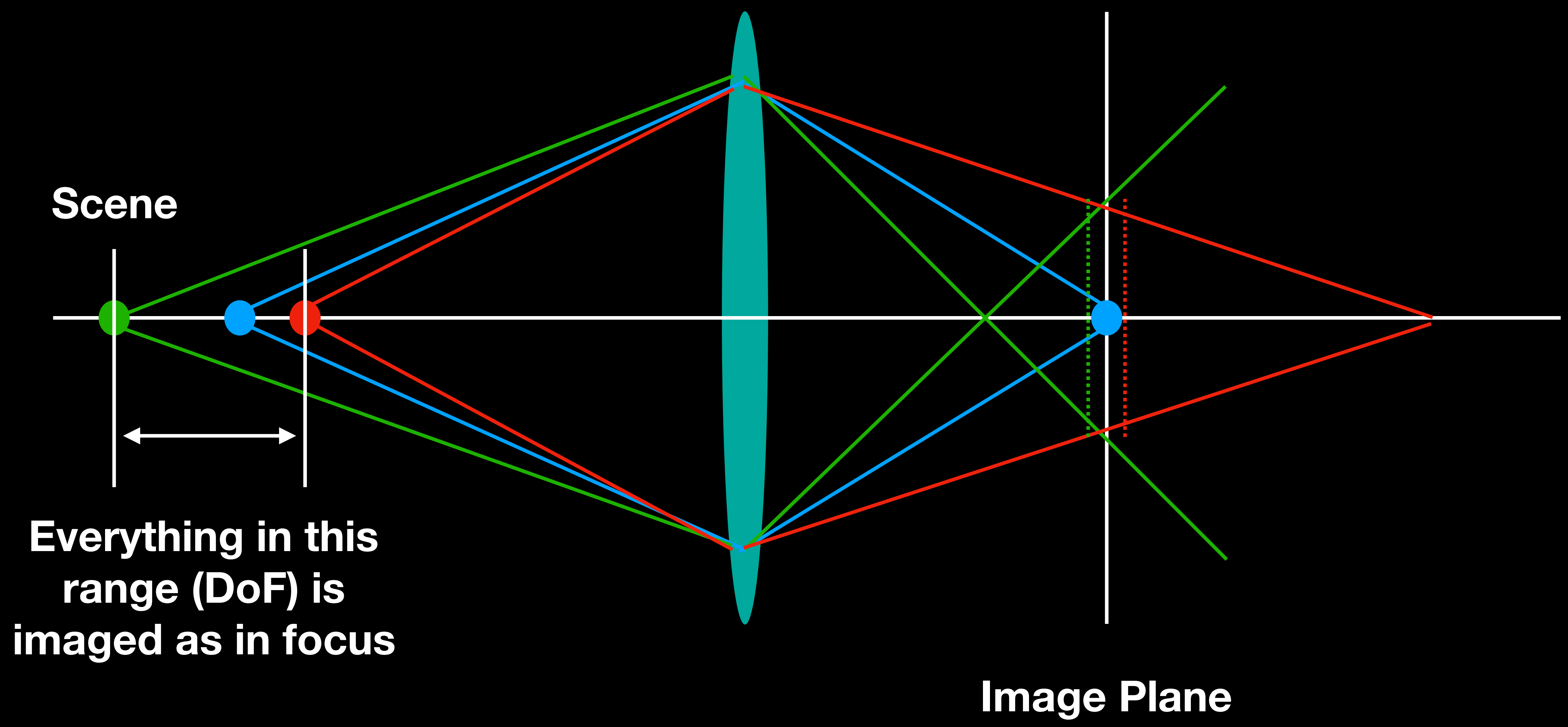


# Optics - Image Formation

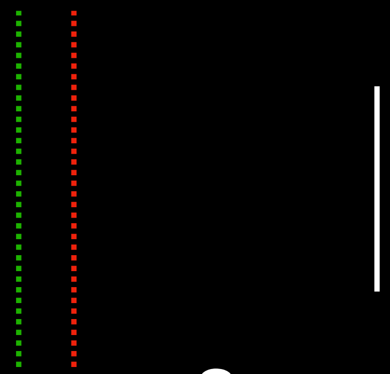


**Suppose the image sensor was square and had the following height**

# Optics - Image Formation

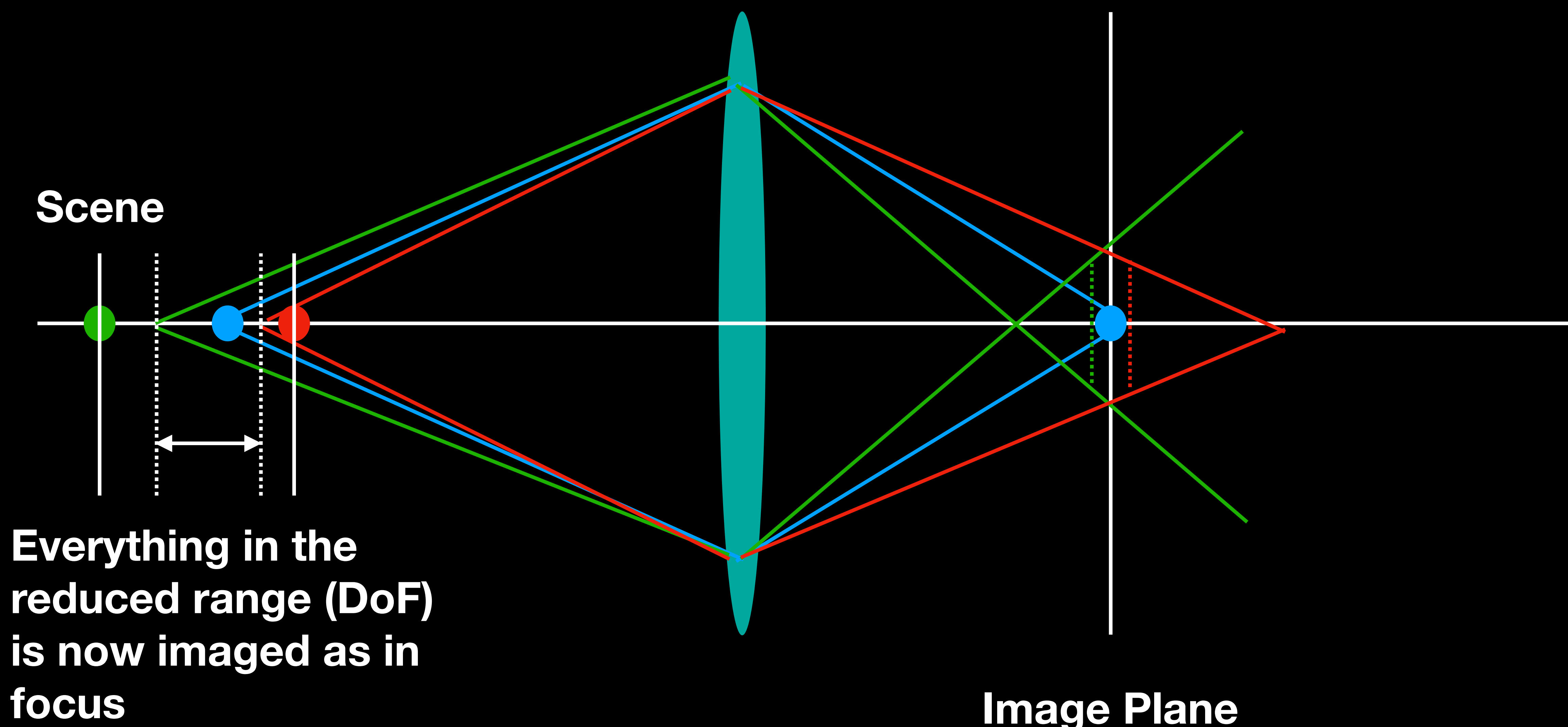


# Optics - Image Formation

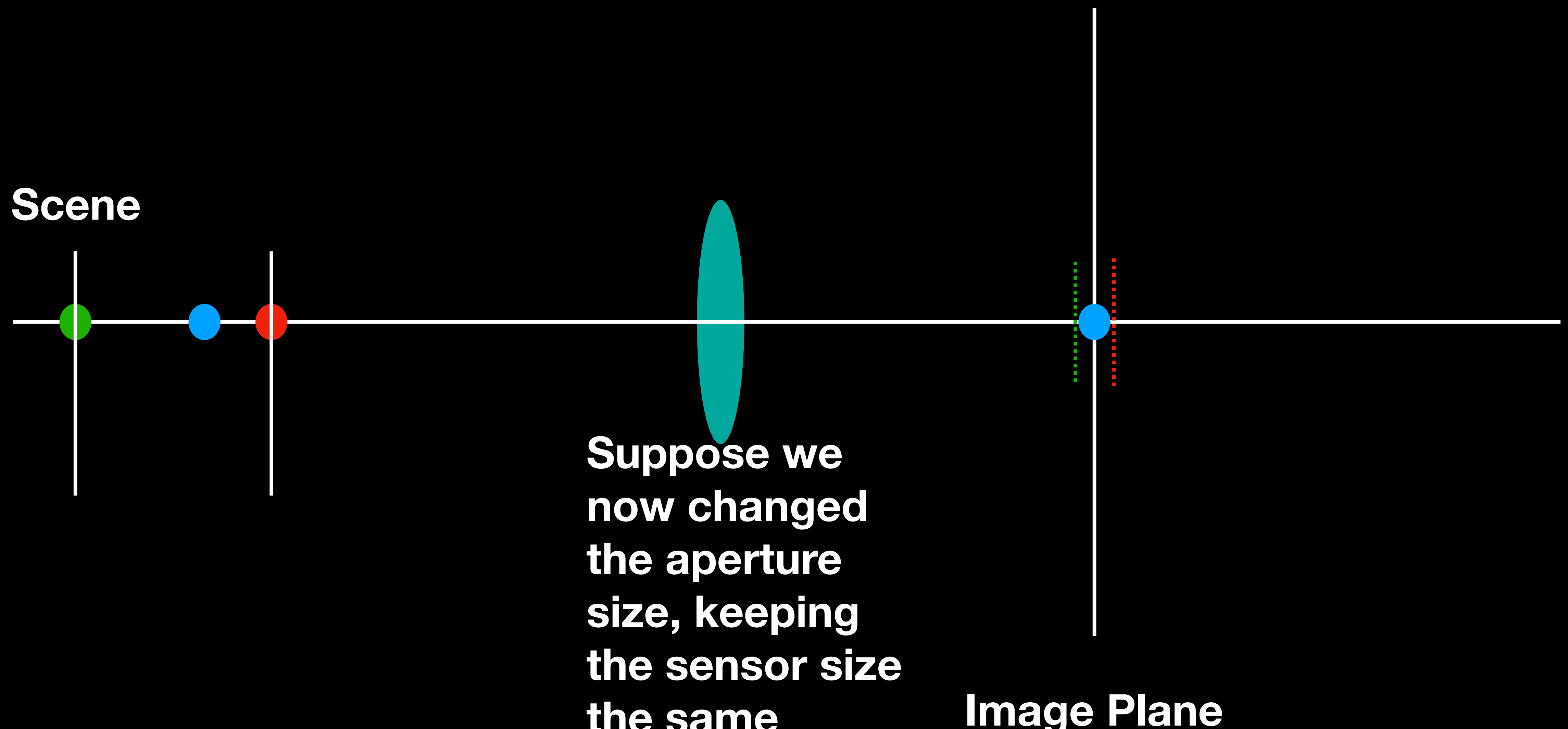


**Suppose the image sensor was square and NOW had the following reduced height**

# Optics - Image Formation

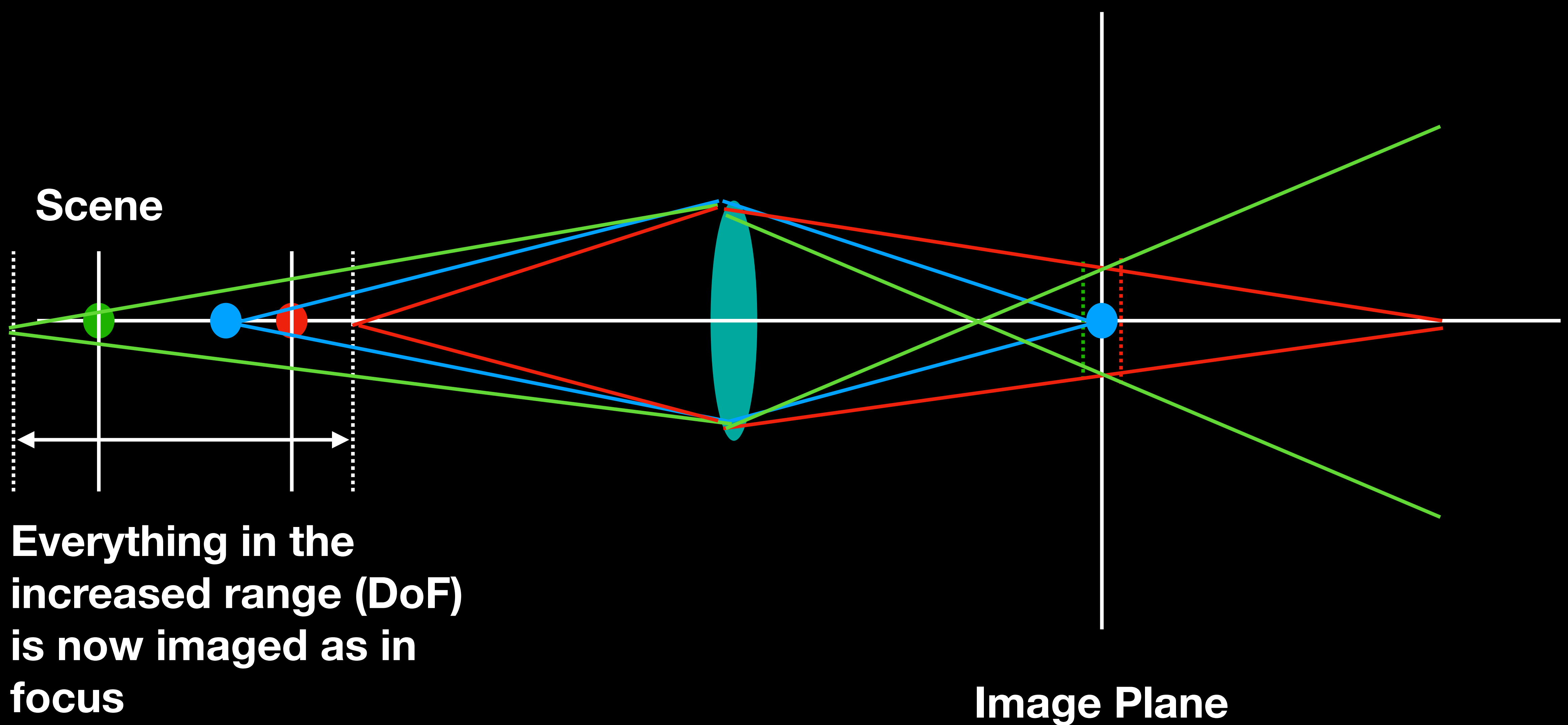


# Optics - Image Formation




Suppose we now changed the aperture size, keeping the sensor size the same

# Optics - Image Formation




Everything in the increased range (DoF) is now imaged as in focus

# Optics - Image Formation

- By construction, there will always be a distance in the scene plane (=focus position) that will be in focus in the image plane.

- The size of the region around the focus position that will be in focus in the image plane increases as the aperture decreases and / or the pixel size increases.

- The size of the region around the focus position that will be in focus in the image plane decreases as the aperture increases and / or the pixel size decreases.

- What type of effects result from changing the DoF?

# Optics - Image Formation

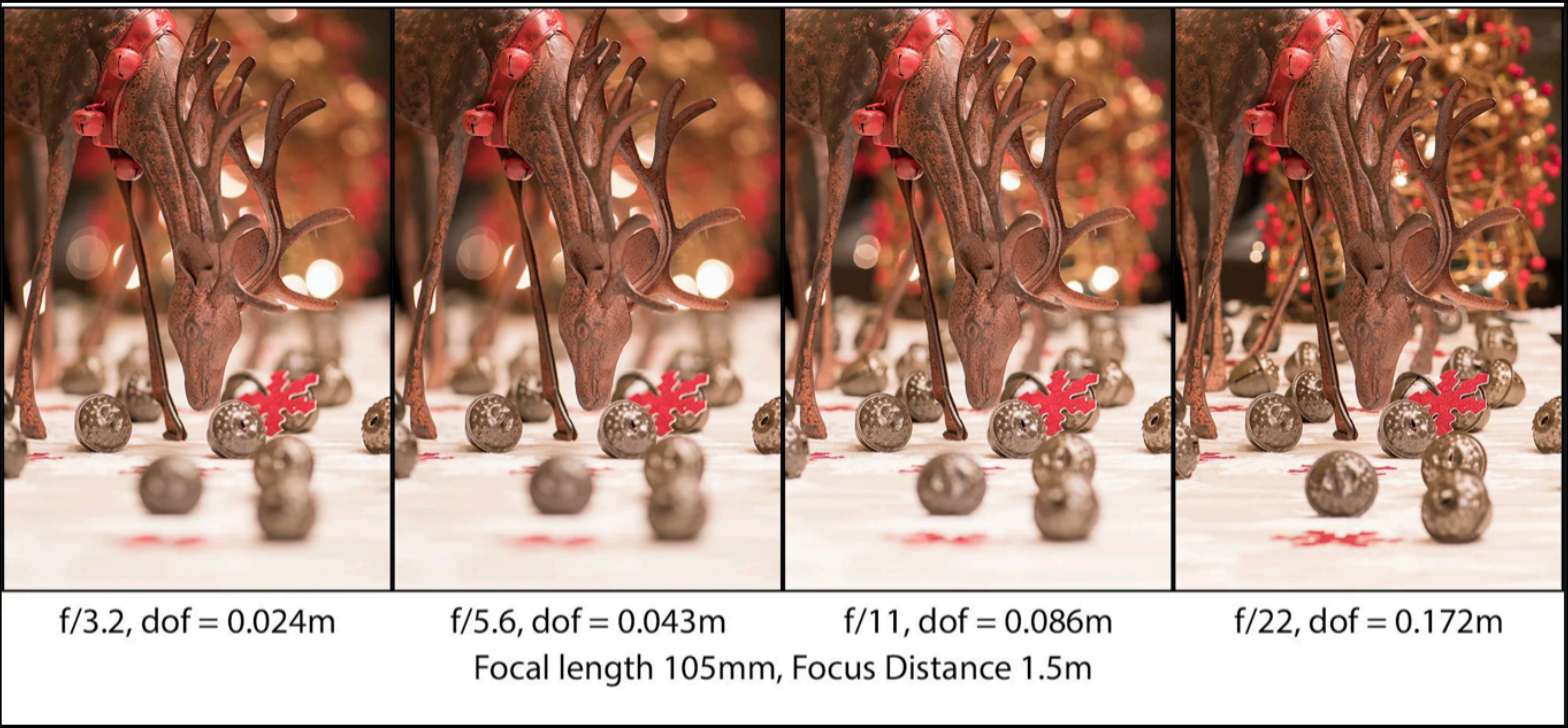



**Image courtesy of Photography Life**

**Here, the aperture is changed.**

**The left most image has a small / shallow DoF.**

**The right most image has a large / deep DoF.**

# Optics - Image Formation

Regarding resolution:

- As the pixels become smaller and more dense, the resolving power / resolution increases.
- As the pixels become larger (think pixel binning), the resolving power / resolution decreases.
- We will now examine resolving power through the modulation transfer function (MTF).

# Optics - Lenses

There is a wide range of lens options:

- Macro, wide angle, telephoto
- Apertures:  f1.4, f1.8, f2.0, f2.8 … f16
- Single element to N lens elements
- Concave and convex lenses
- Plastic, glass and associated lens coatings, etc.
- Telescopic / folded optics, etc.

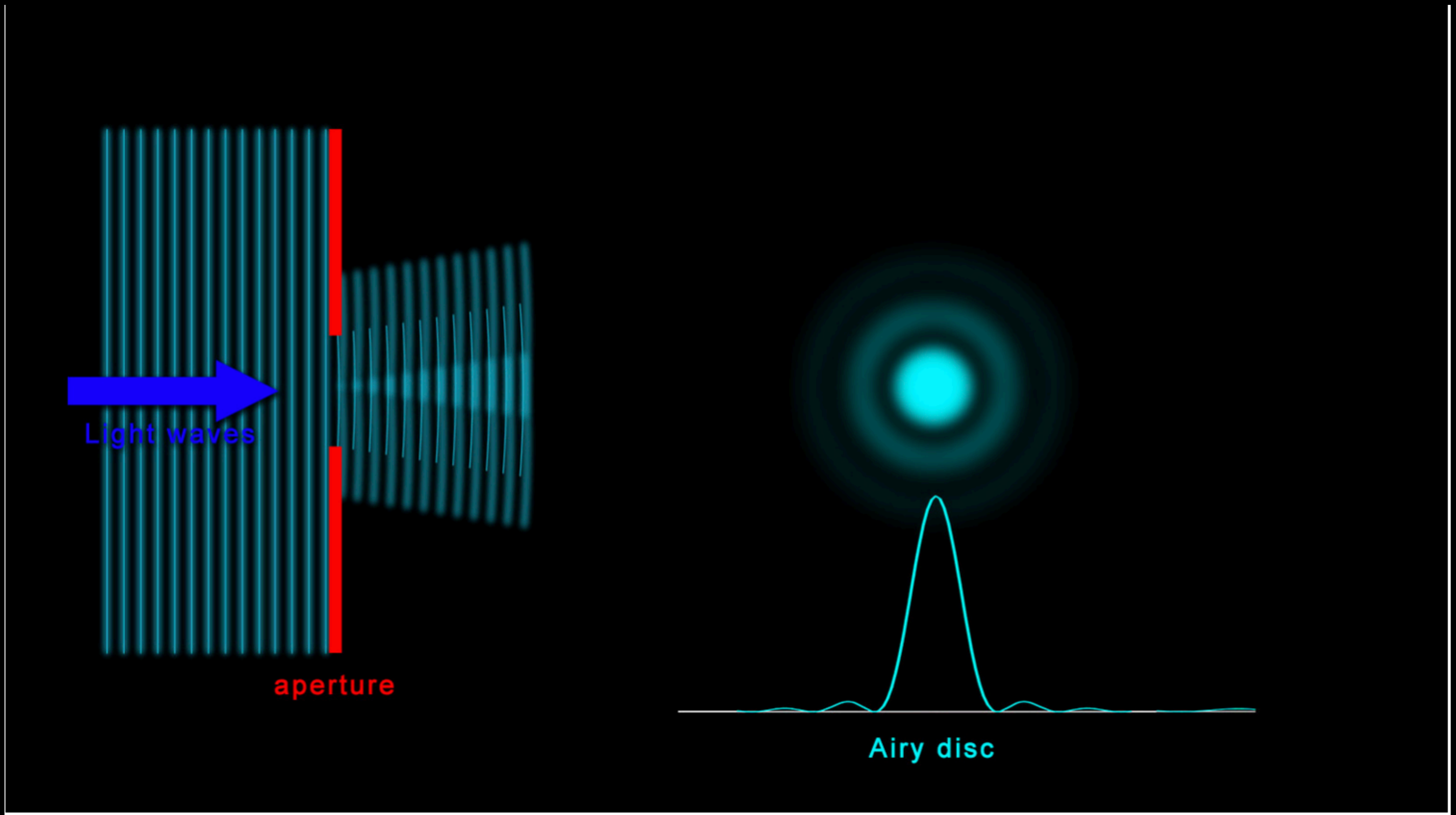
# Optics - Lenses

- As stated previously, the purpose of the lens is to gather light and to focus light.
- Common to each lens, is the concept of DoF and the resulting image.
- Manufacturing costs, quality of output, and use cases all serve as lens design parameters.

# Optics - Design

- Because every lens must have a finite aperture (diameter), diffraction is a common, physical property.
- From the wave theory of physics, diffraction results in (downstream) constructive and destructive interference.
- For circular apertures, this interference is physically measured (by an imaging sensor) as Airy disk.
- The diameter of the Airy disk is given by: $d_{Airy} = 2.44\lambda N$ , where N is the f-number of the lens.

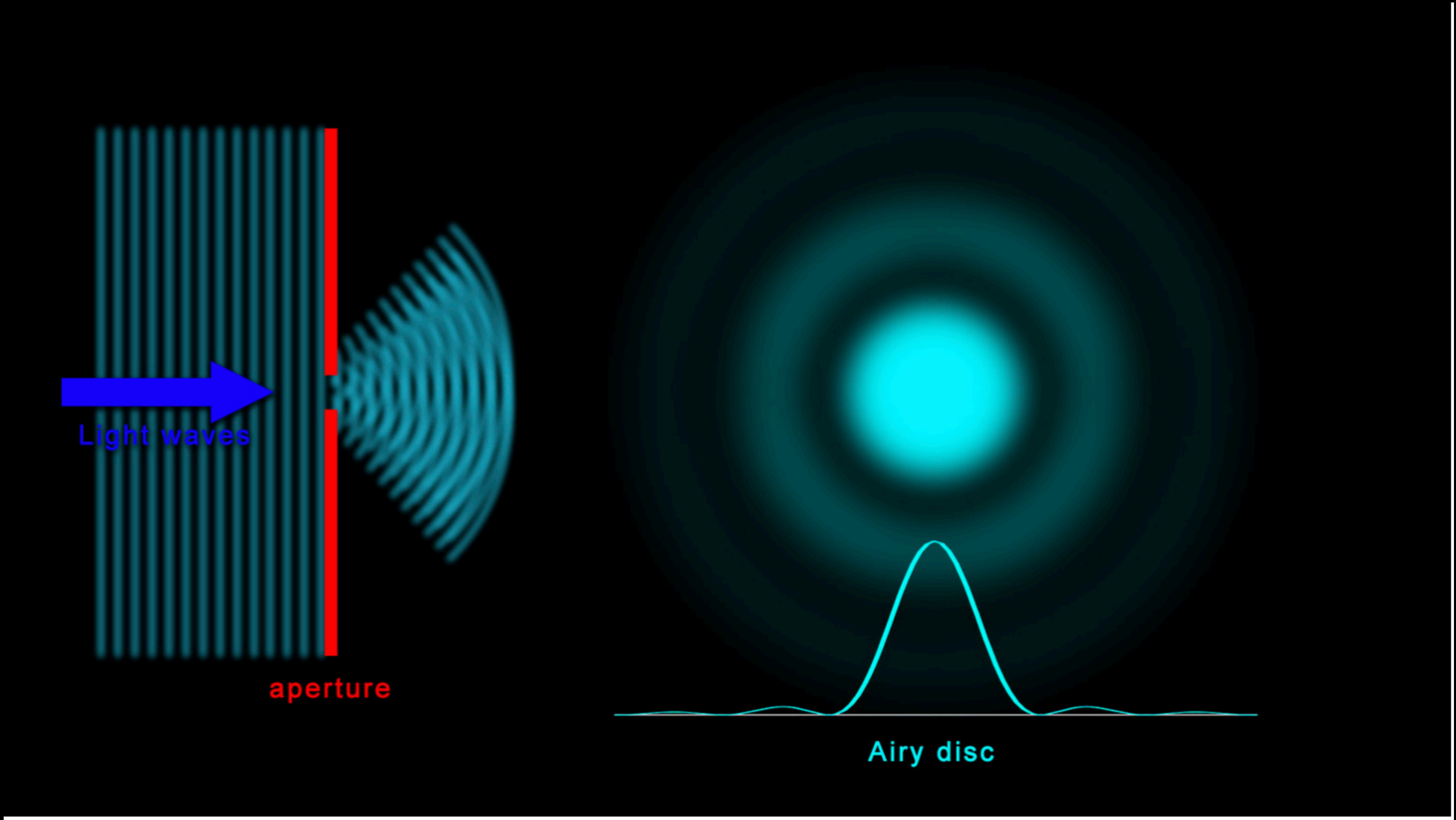
# Optics -Design




A large lens opening will make the Airy disk small.

**Image courtesy of FStoppers**

# Optics - Design




If the lens opening becomes smaller, the Airy disk will grow larger.

**Image courtesy of FStoppers**

# Optics - Design

- To minimize diffraction (and subsequently, the diameter of the airy disk) a very large aperture is needed.

Technically:

- A Bessel function output results, when light is diffracted through a circular aperture.

- The square of the Bessel function (=airy disk) is then captured by the image sensor (=point spread function (PSF))

- When the fourier transform of PSF (=airy disk) is computed, the optical transfer function (OTF) is obtained.

- When we the magnitude of the (OTF) is computed, the modulation transfer function (MTF) is obtained.

The MTF describes how contrast decreases as a function of spatial frequency (line pairs /mm).

The perfect MTF curve would be a straight line with contrast of 1 (or modulation = 100 percent), for the full range of spatial frequencies being evaluated.

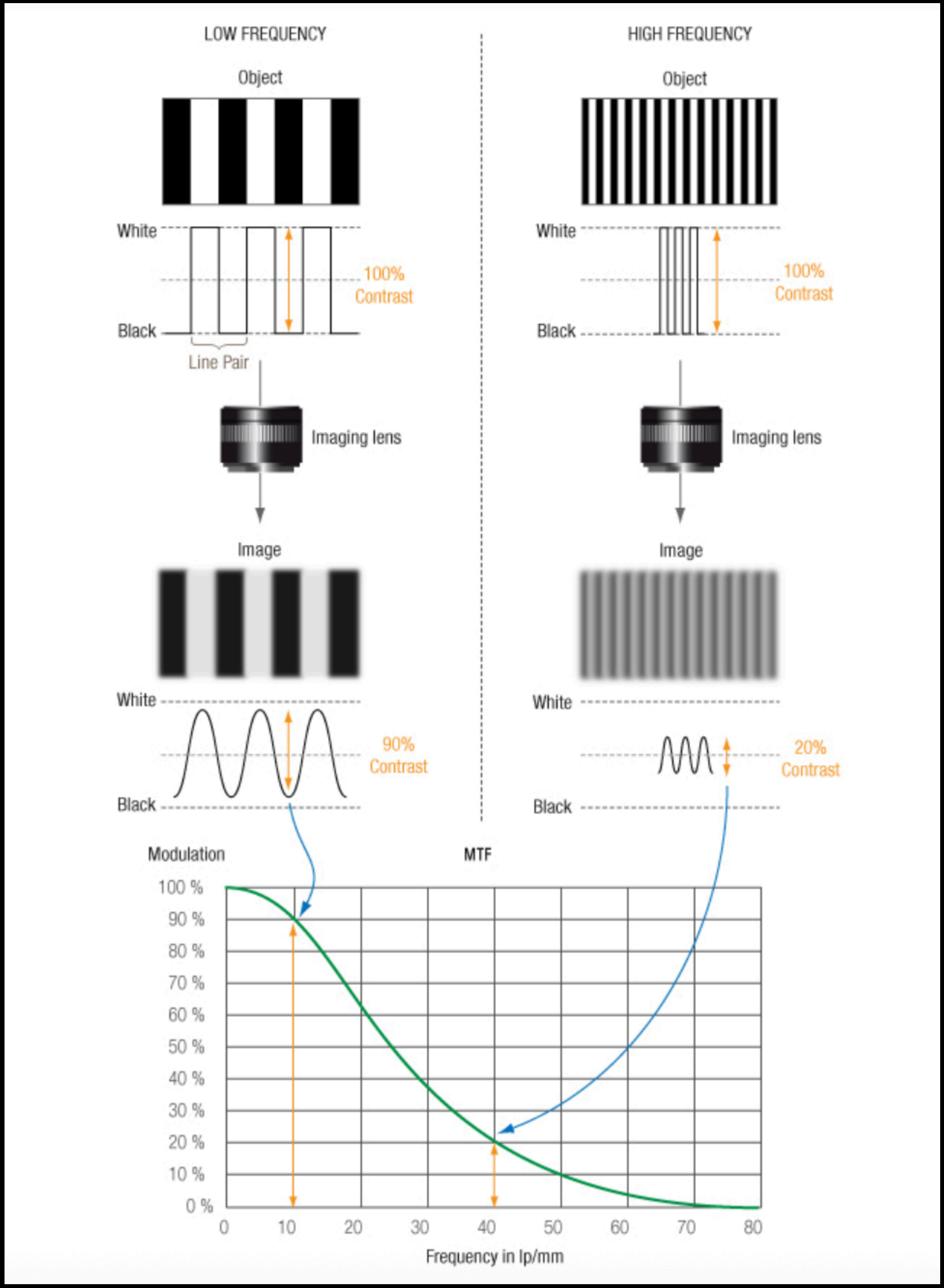



Image courtesy of DXOMark

# Optics - Design

- The pixel dimension also has an associated MTF.
- Consider a square pixel.
- The square pixel integrates light over a given area.
- Mathematically, the pixel performs a convolution operation with the scene.
- $g(x, y) = I(x, y) * h_{pix}(x, y)$ where $h_{pix}(x, y) = \frac{1}{a * a} rect(\frac{x}{a}) rect(\frac{y}{a})$

# Optics - Design

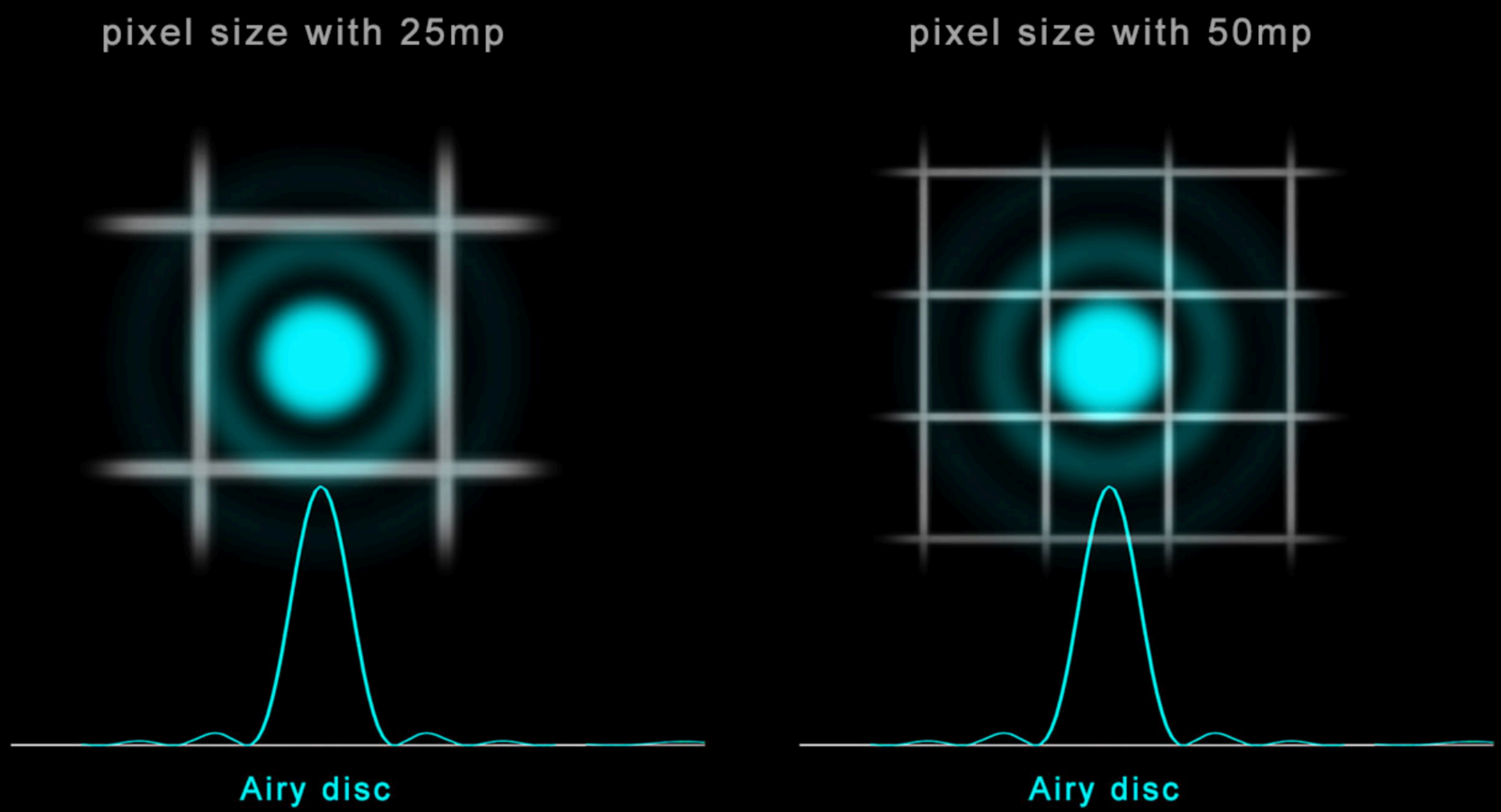
- We compute the Fourier transform of $h_{pix}(x,y) = \frac{1}{a*a} rect(\frac{x}{a}) rect(\frac{y}{a})$ obtaining the OTF: $H_{pix}(f_x, f_y) = sinc(af_x) sinc(af_y)$

- The MTF is given by: $|H_{pix}(f_x, f_y)| = |sinc(af_x) sinc(af_y)|$

- The system MTF is now given by: $MTF_{system} = MTF_{diffraction} MTF_{pixel}$

- The picture on the next slide, illustrate the physical interaction between the two quantities in the spatial domain.

# Optics - Design




**Resolving power limited by pixel size.**

**Resolving power limited by aperture size (diffraction).**

**Image courtesy of FStoppers**

# Optics - Design

- There are several important “numbers” associated with $MTF_{diffraction}$ $and$ $MTF_{pixel}$
- For $MTF_{diffraction}$ the spatial frequency at which MTF = 0 is given as $f_{cutoff} = \frac{1}{\lambda N}$
- When we stop down a lens, the cutoff frequency decrease.
- The Nyquist frequency for $MTF_{pixel}$ is given as $f_N = \frac{1}{2a}$
- At the Nyquist frequency, $MTF_{pixel} \sim 0.64$

# Optics - Design

- In terms of the system MTF, we observe that we have a diffusion MTF curve and a pixel MTF curve.

- Both curves decrease, as the spatial frequency increases.

- To obtain the best system MTF curve (=the product of the two curves), we would then need to:

1) use the largest aperture possible (to minimize the airy disk diameter / minimize the PSF due to diffraction)

2) use the smallest pixel dimension possible (to maximize the pixel MTF)

# Optics - Design

Things to consider:

- Larger apertures require large lenses.
- Large lenses weigh more, cost more, and are more difficult to manufacture.
- Additional degradations (like lens aberrations) could now also become potentially significant, introducing yet another MTF term.
- Smaller pixels mean less light gathering capabilities.
- Less light gathering capabilities means higher image noise.

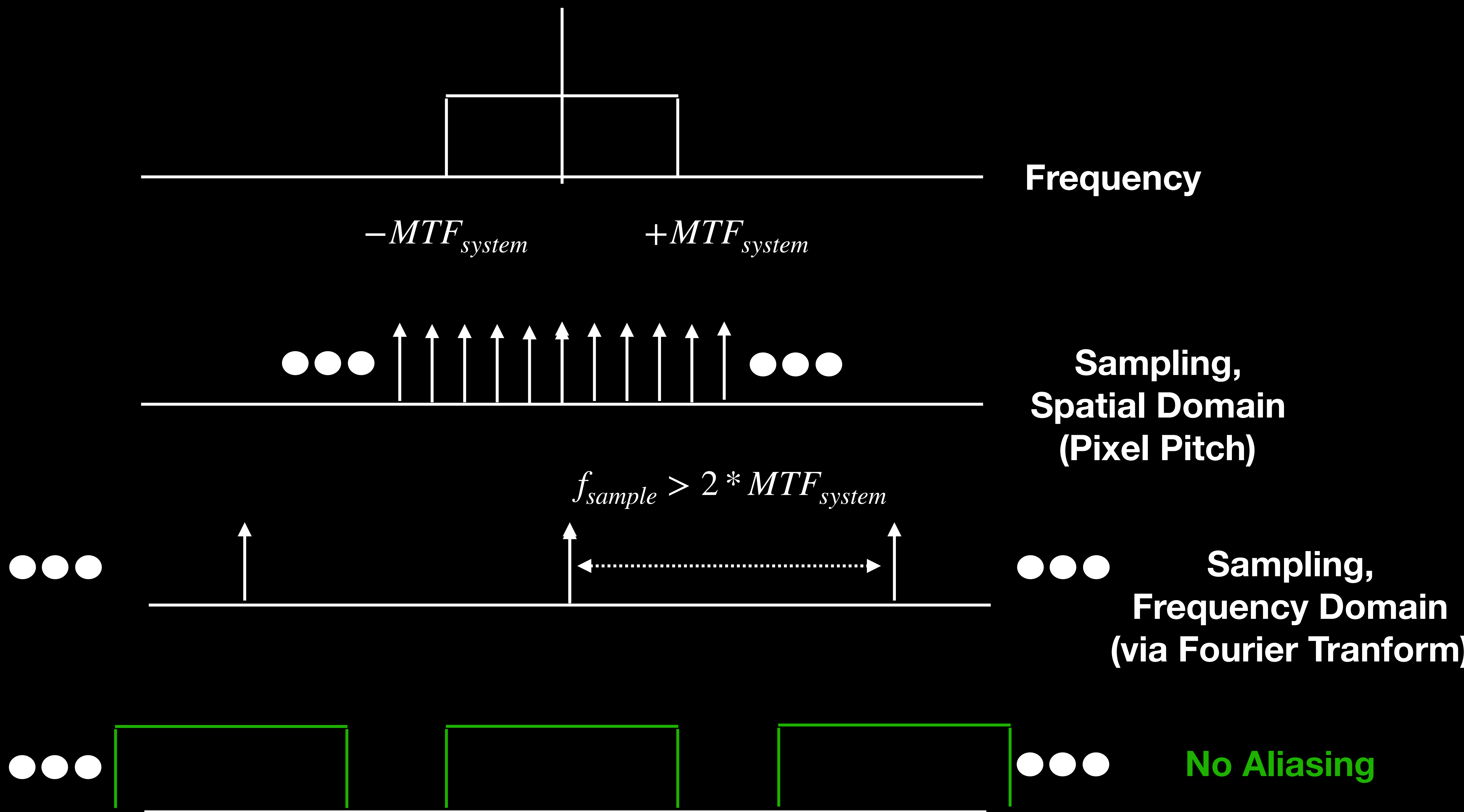
# Optics - Design

- Fortunately, before trying to maximize the system MTF, there is one additional factor to consider.
- This factor is sample frequency, and the concept of aliasing.
- From Nyquist sampling theory, to prevent aliasing artifacts, the sampling frequency needs to be twice the frequency bandwidth of the baseband signal.
- But where is the sampling coming from, in our current system?

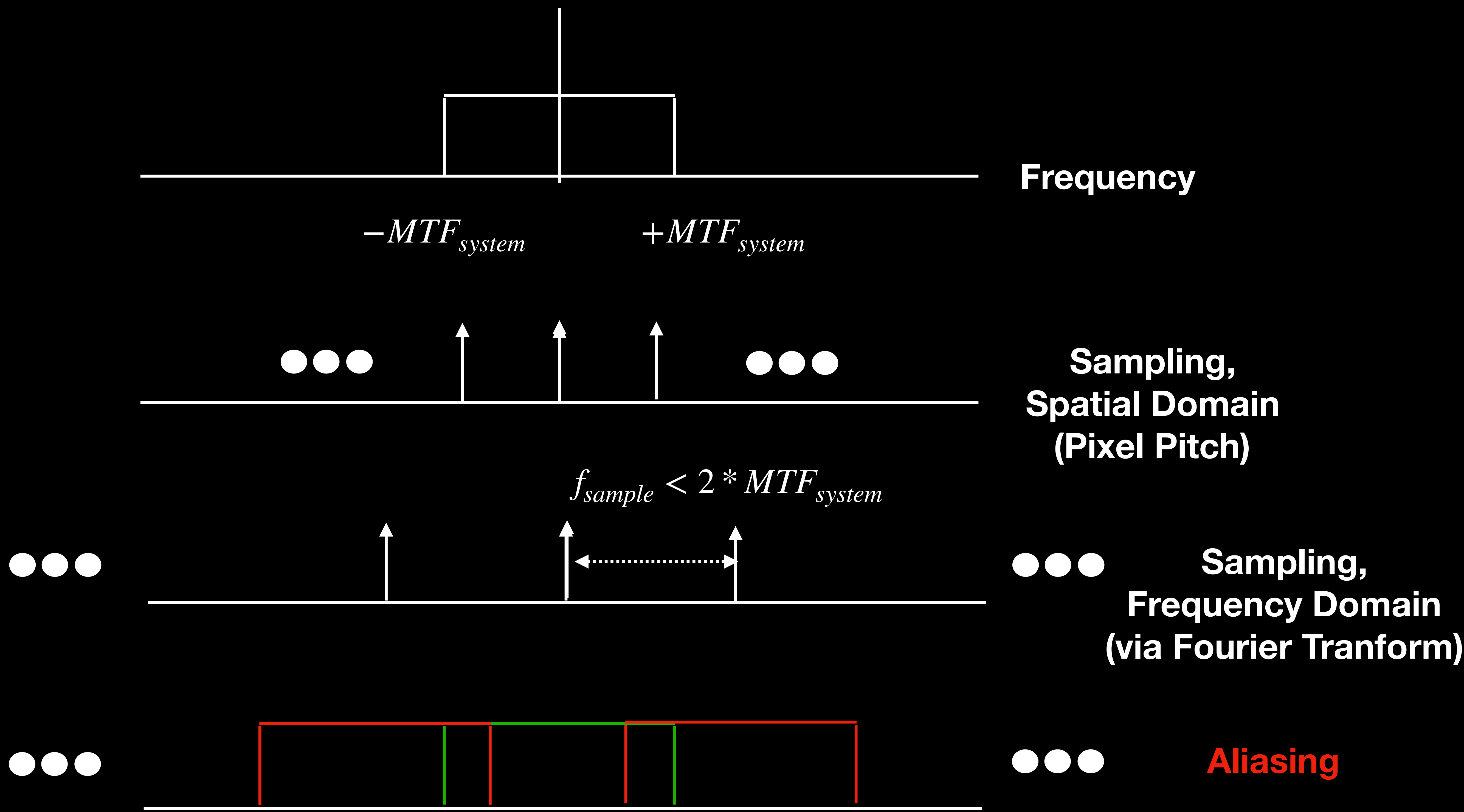
# Optics - Design

- The spatial sampling comes from the pixel pitch.
- The fourier transform of the pixel pitch sampling defines the sampling frequency used in Nyquist’s Theorem.
- In practice, the pixel pitch is actually 2x larger than the distance between adjacent pixels, because a bayer color filter array (CFA) is used.
- This increases the likelihood of aliasing.
- With this in mind, let’s illustrate the no aliasing, aliasing, and no aliasing nyquist rate cases for the Nyquist sampling theorem.

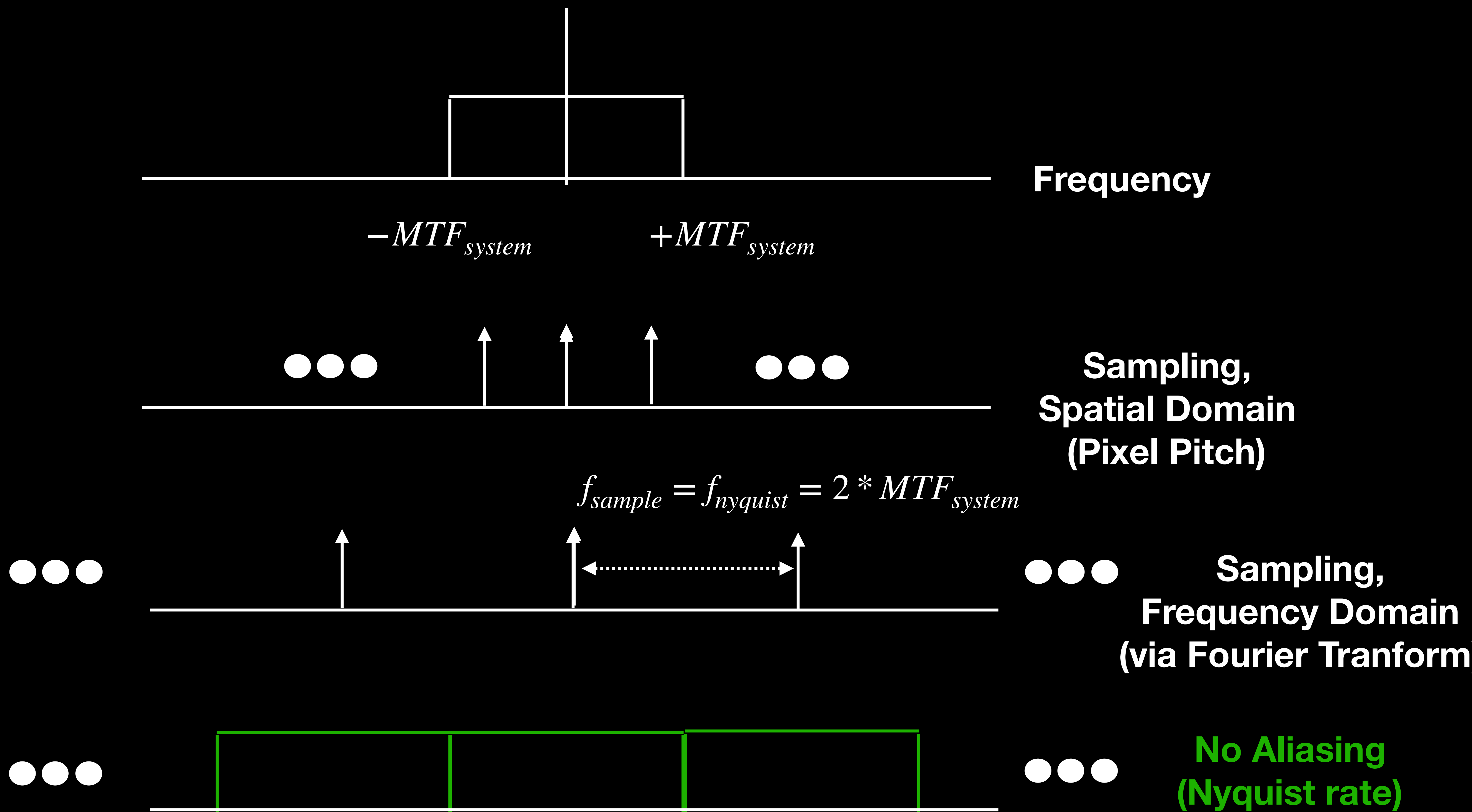
# Optics - Design



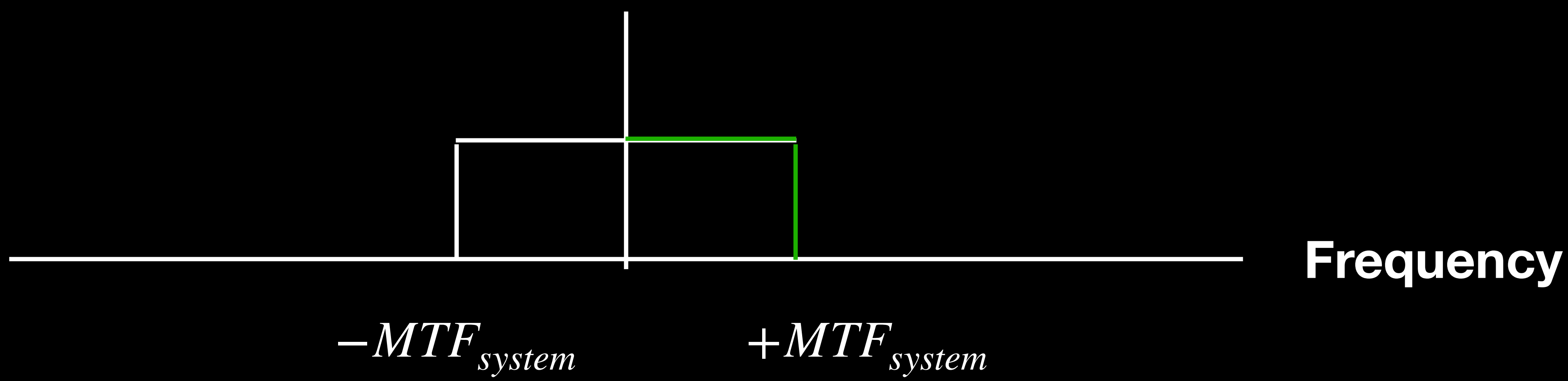
# Optics - Design



# Optics - Design



# Optics - Design



# Optics - Design

- Determine the bandwidth needed, to satisfy the nyquist criteria to prevent aliasing.

- i.e. Determine the sampling frequency, from the pixel pitch.

- Design the MTF attributed to diffraction and the MTF attributed to pixel size, so that the system MTF satisfies the nyquist criteria for all use cases.

- i.e. Every use case supported - different aperture stops and different pixel sizes due to binning / non-binning - need to be considered.

# Summary

- Lenses are used, to gather / collect light.
- With every lens, diffraction will occur, due a finite lens aperture.
- Lenses with smaller apertures will have larger airy disks, and hence, larger PSF’s.
- With every lens, different DoF’s will result, based on the focal length of the lens, the focus position of the lens, the pixel size of the sensor, and the size of the lens aperture.
- This produces a wide variety of images, with different degrees of blur around the lens focus position.

# Summary

- Different lenses are designed, to address different use cases.

- The MTF curve captures the relationship between contrast and spatial frequency.

- MTF curves exist for the lens and the sensor.

- To prevent or minimize aliasing, the system MTF needs to be small, beyond the Nyquist frequency.

- The Nyquist frequency is defined by the sensor pixel pitch.