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

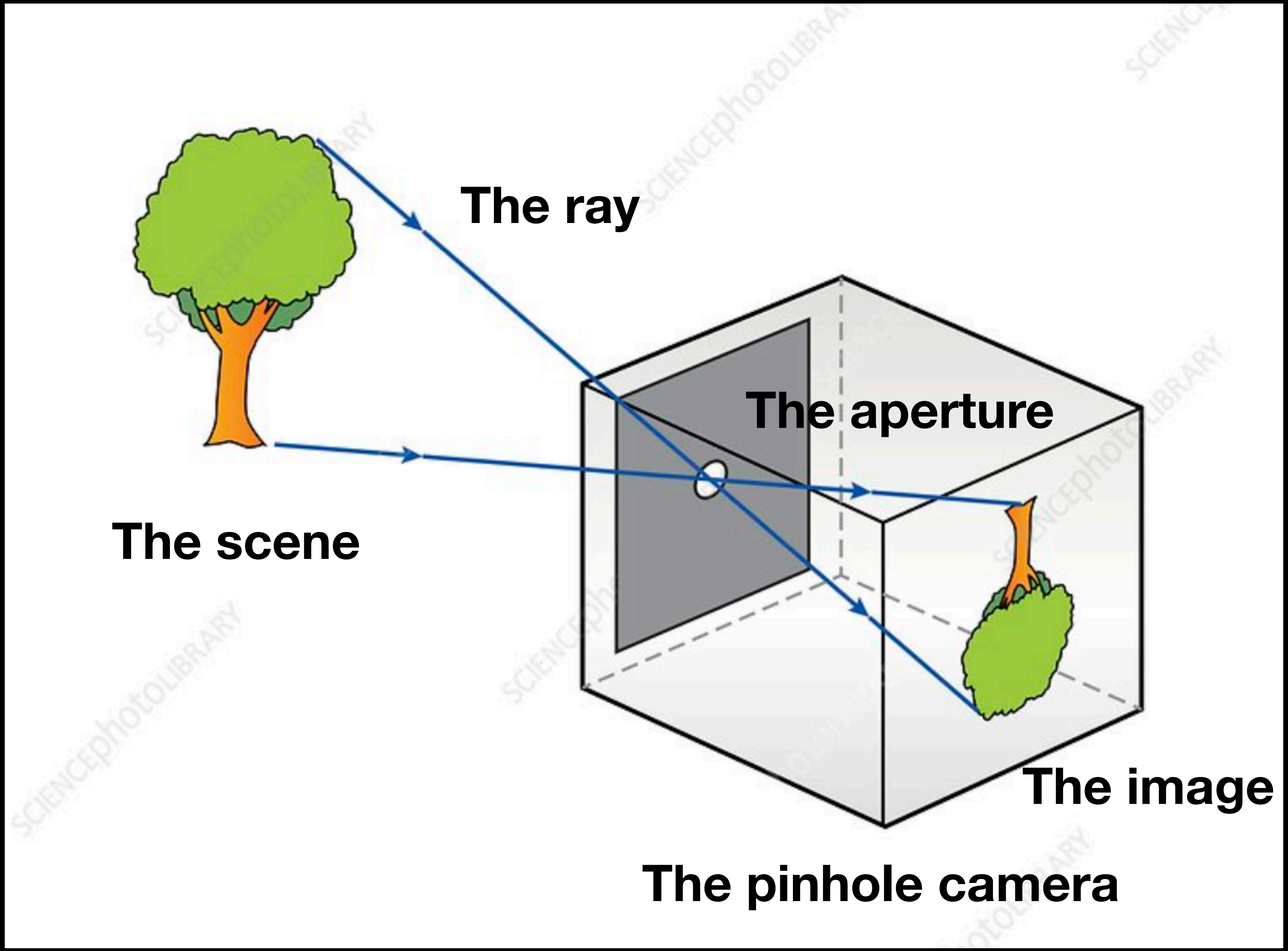


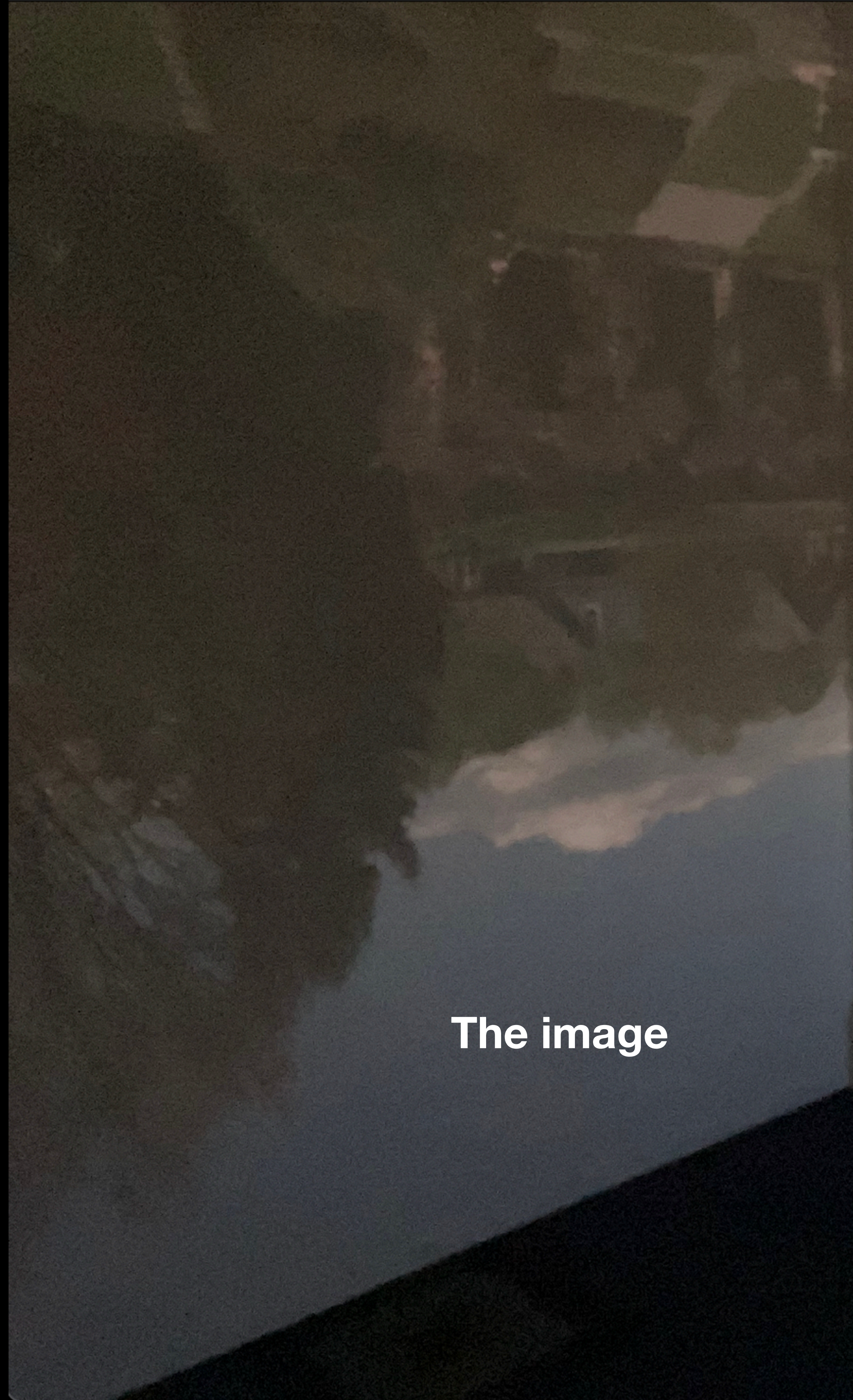
Image courtesy of Science Photo Library

Optics - Image Formation

- Using ray tracing, we can map illumination sources in the scene into an 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.

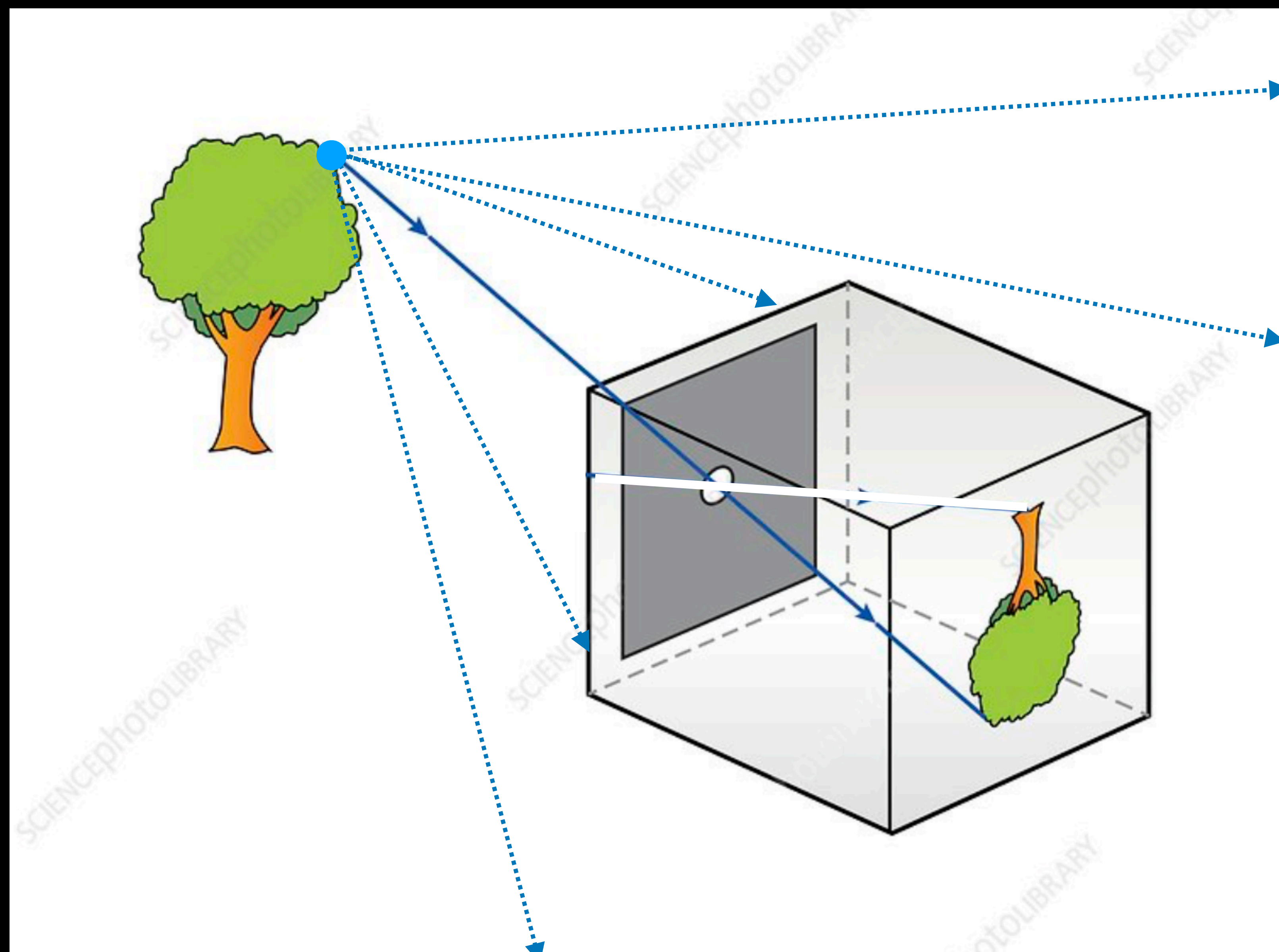

The aperture


The image

Optics - Image Formation

- Are we done?
- Not even close.
- What more is there to do?
- Something about collecting the light?
- This is the whole point of optics / lenses.

Optics - Image Formation



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

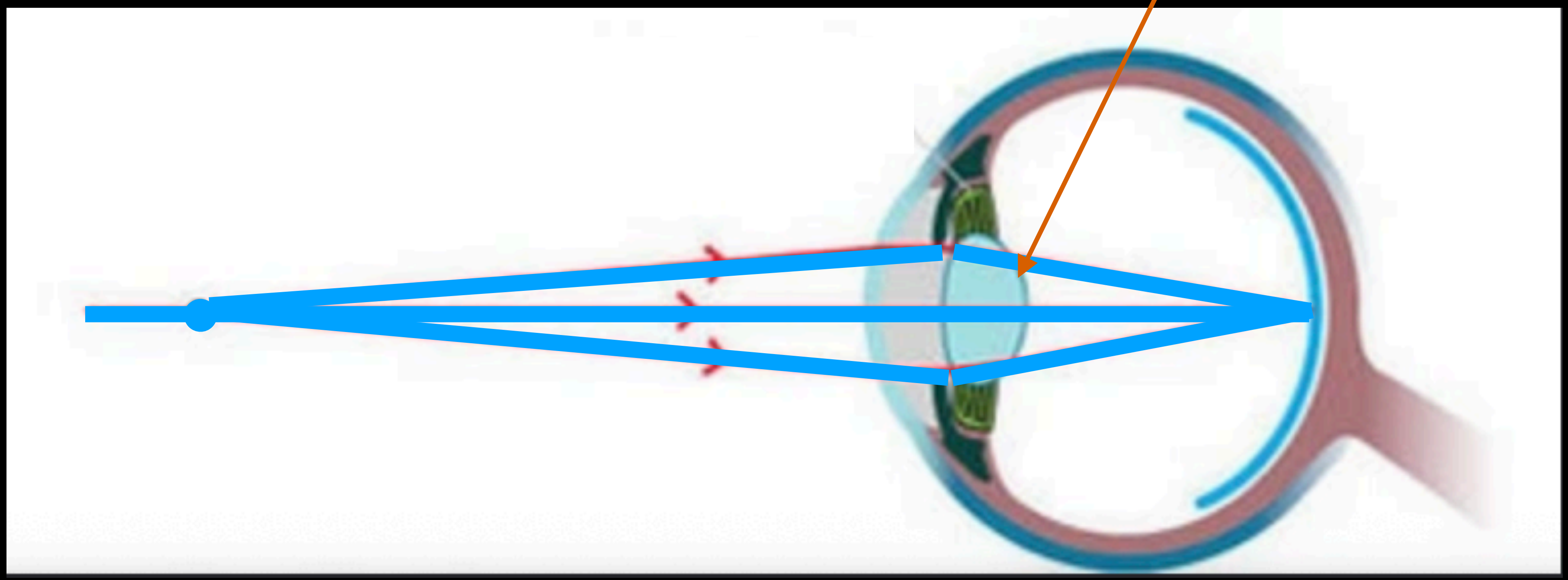
Hence, light gathering becomes a significant issue.

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.

Optics - Image Formation

Lens of the human eye

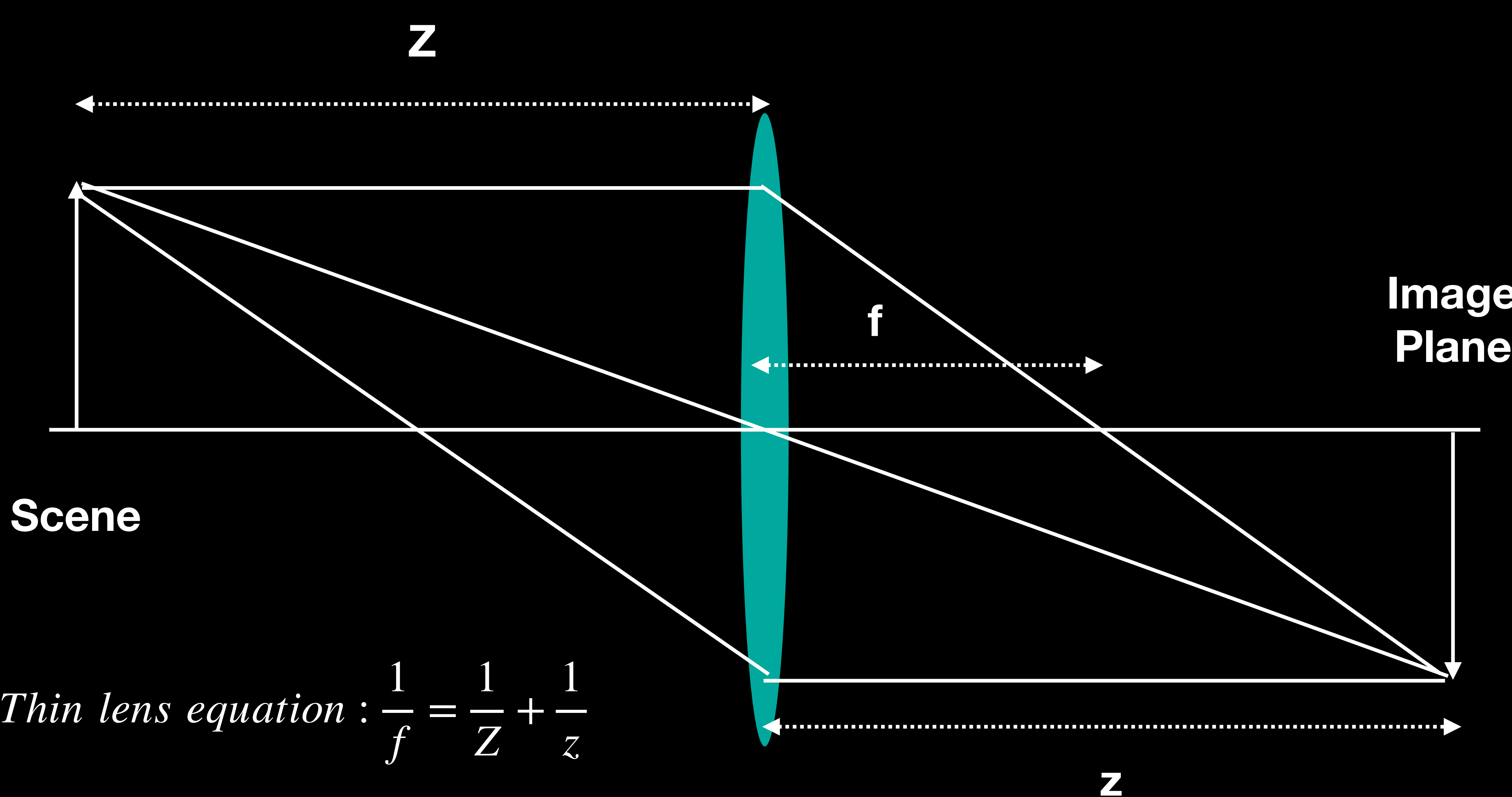


This is how our human visual system does it.

Optics - Image Formation

- From this, we can obtain the most basic lens equation - the thin lens equation.
- The thin lens equation relates objects in the scene to objects in the image plane, via the lens focal length.
- The lens focal length is defined to be the location where parallel rays from the scene (=very distant object point source) focus at a single point.

Optics - Image Formation



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

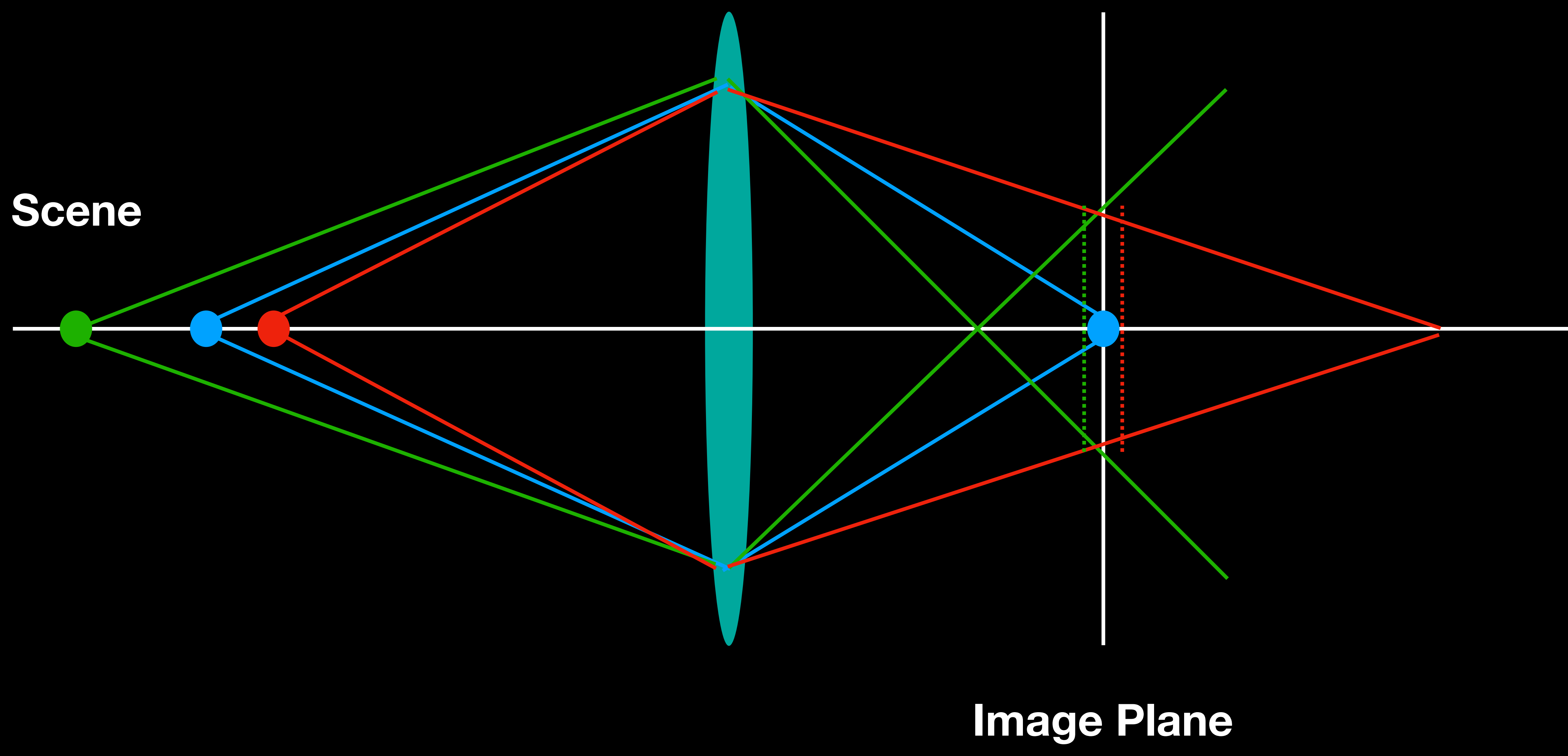
Optics - Image Formation

- 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.
- The depth of field (DoF) determines the “range” of distances in the scene that will be imaged “in focus” on the image plane.

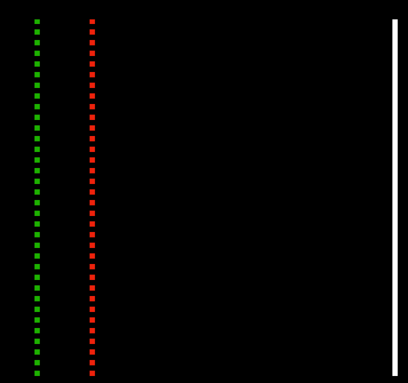
Optics - Image Formation

- The DoF is a function of the focus position.
- i.e. The DoF is larger for distant focus positions and smaller for closer focus positions.
- The DoF is a function of the focal length.
- i.e. The DoF is smaller for larger focal lengths and larger for smaller focal lengths.
- In addition, the DoF is also a function of pixel size and aperture size; we now illustrate both of these cases.

Optics - Image Formation

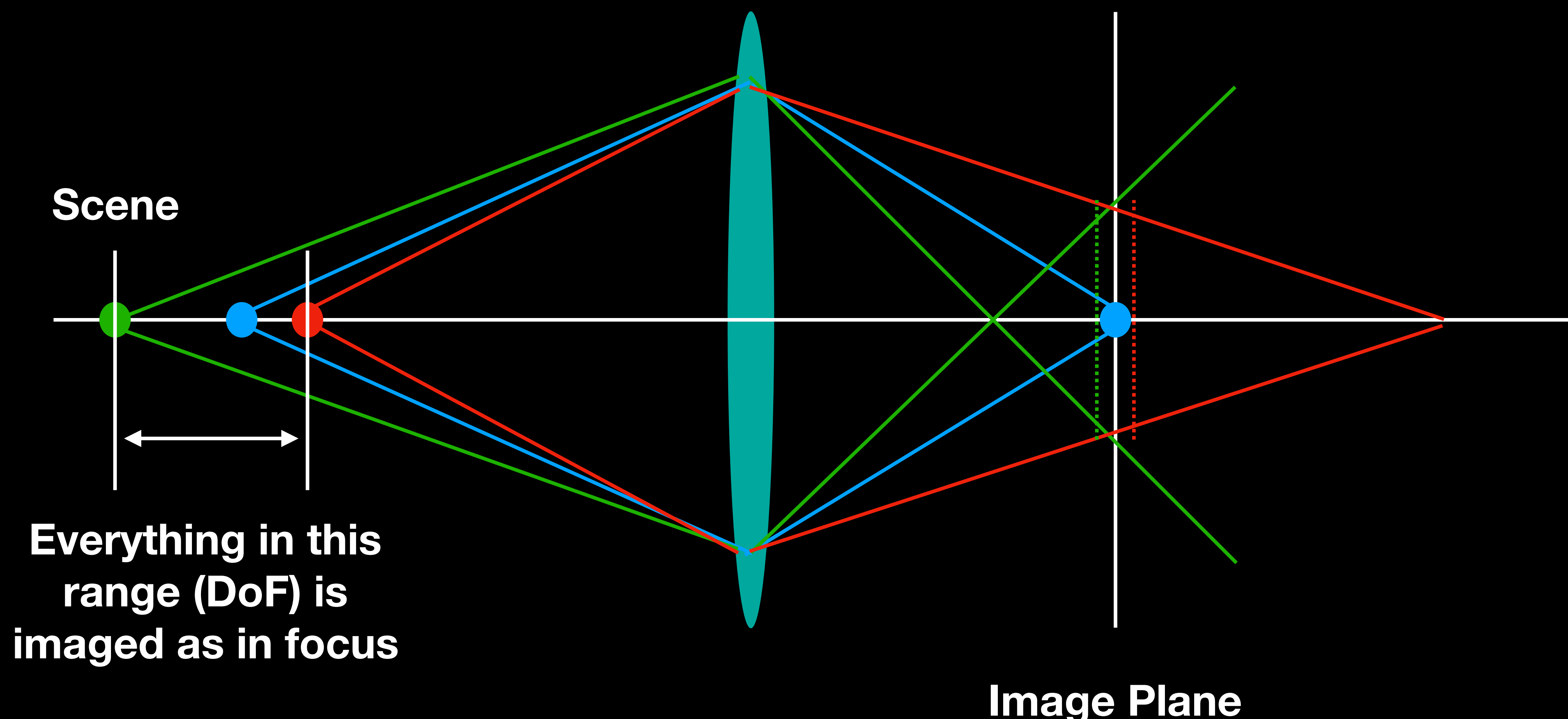


Optics - Image Formation

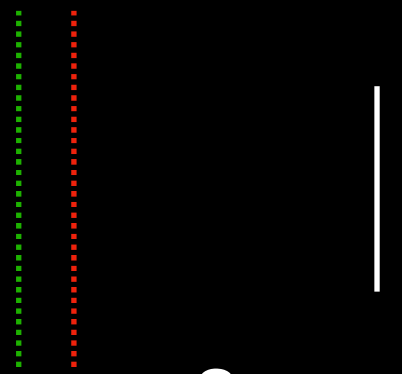


Suppose the image sensor was square and had the following height

Optics - Image Formation

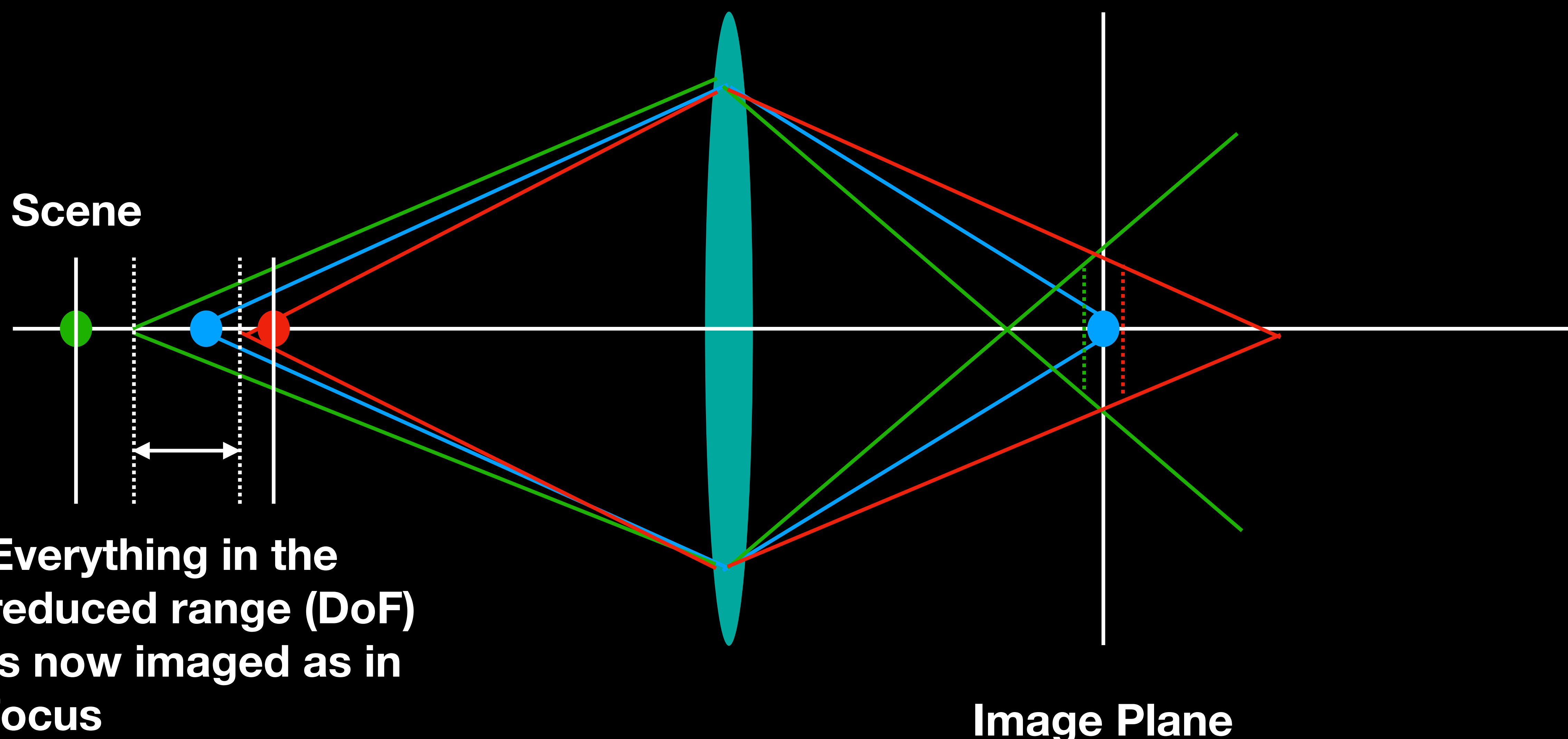


Optics - Image Formation



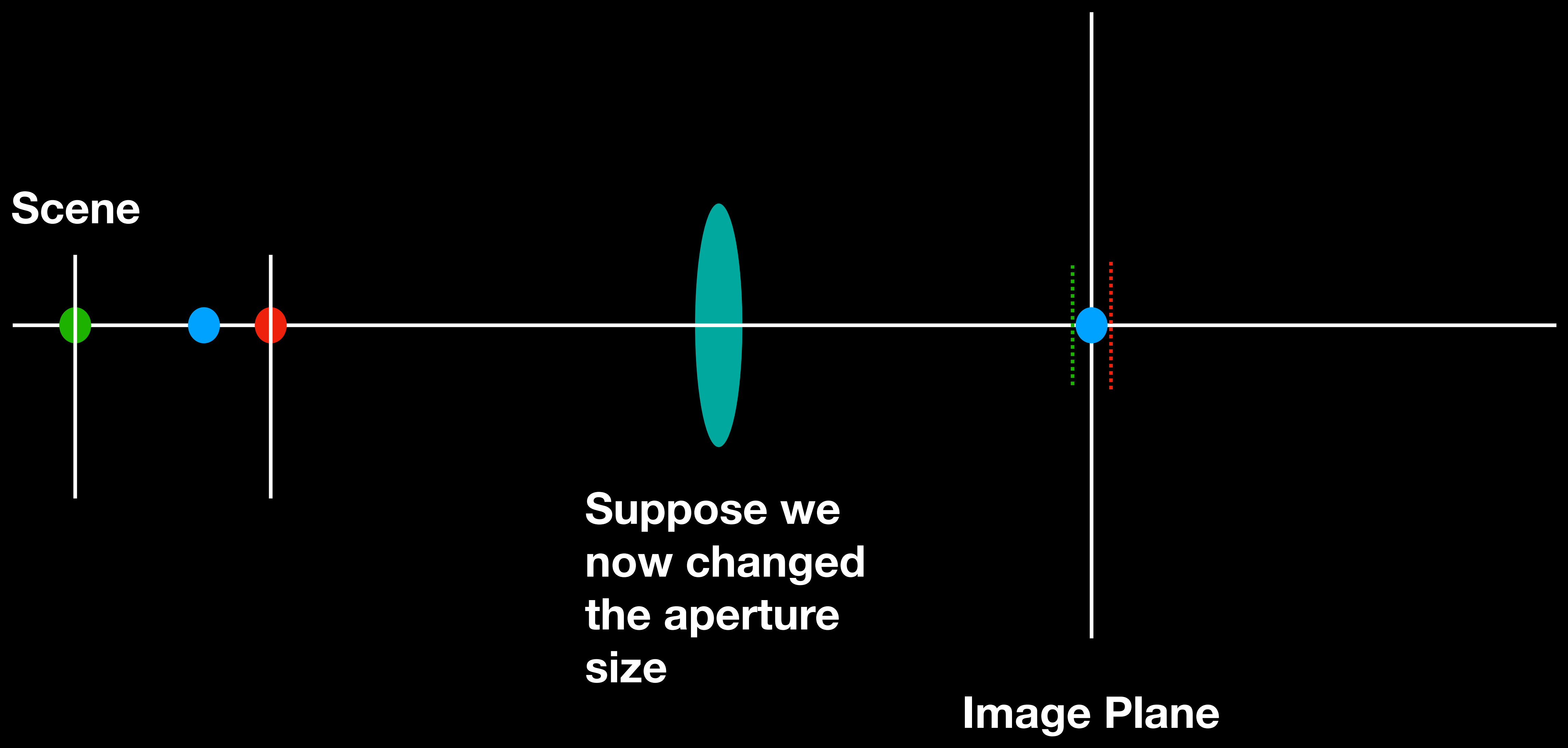
Suppose the image sensor was square and NOW had the following reduced height

Optics - Image Formation

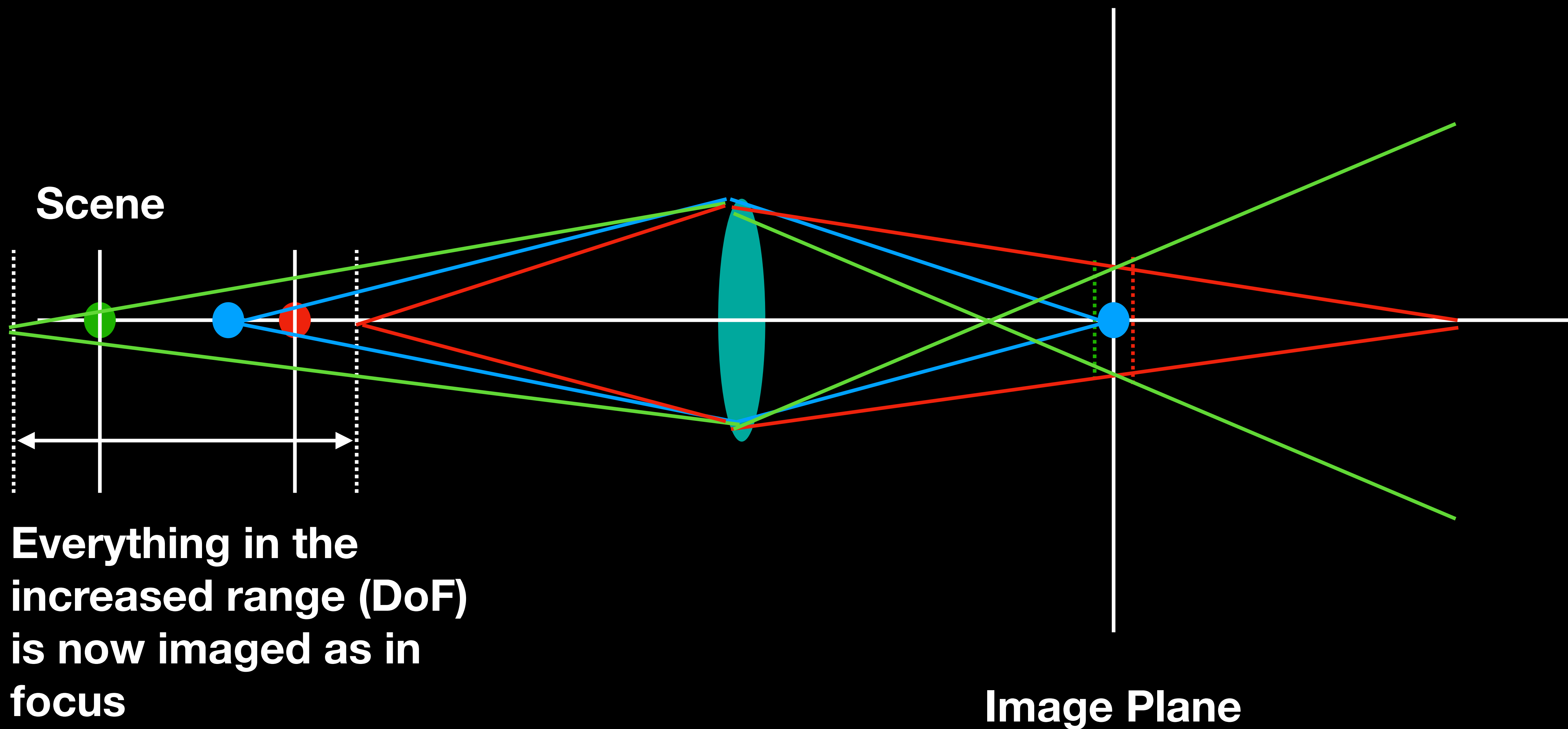


Everything in the reduced range (DoF) is now imaged as in focus

Optics - Image Formation



Optics - Image Formation



Optics - Image Formation

- There will always be a distance in the scene plane that will be in focus in the image plane.
- The size of the region around the scene plane 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 scene plane that will be in focus in the image plane decreases as the aperture increases and / or the pixel size decreases.

Optics - Image Formation

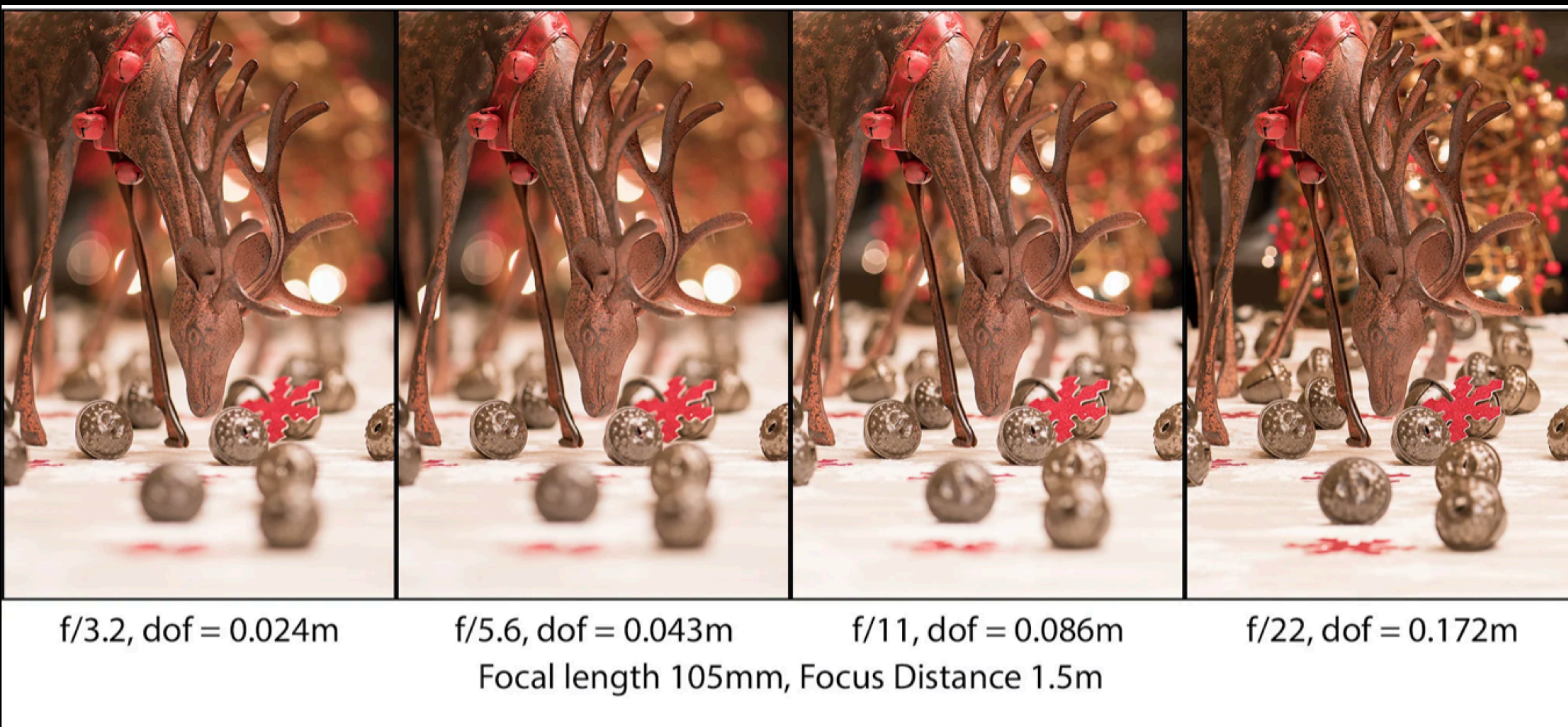


Image courtesy of Photography Life
Here, the aperture is changed.

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 options:

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

Optics - Design

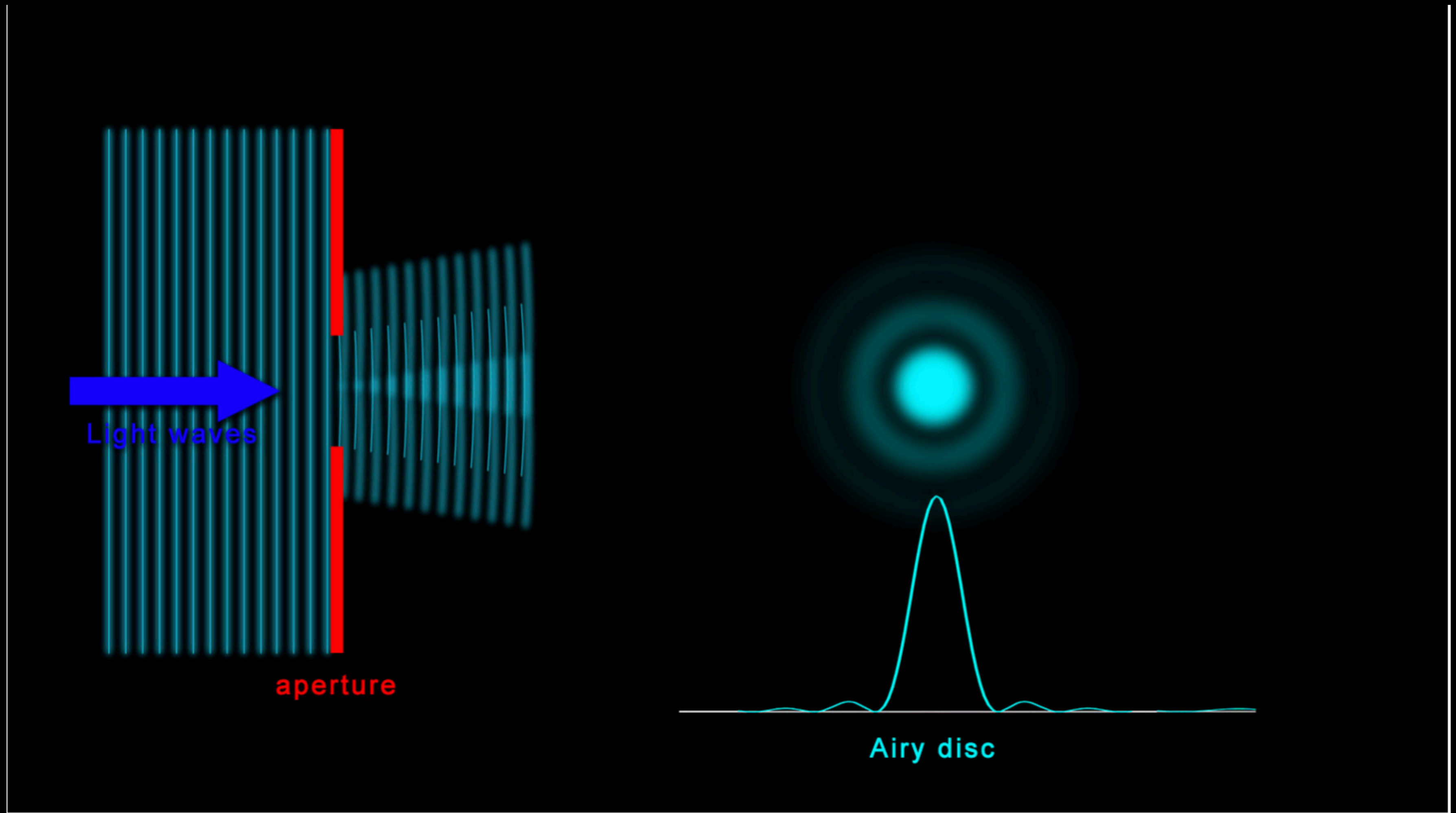
- Because every lens must have a finite aperture (diameter), diffraction is a common, physical property of interest.

- 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 the Airy disk is given by: $d_{Airy} = 2.44\lambda N$, where N is the f-number of the lens.

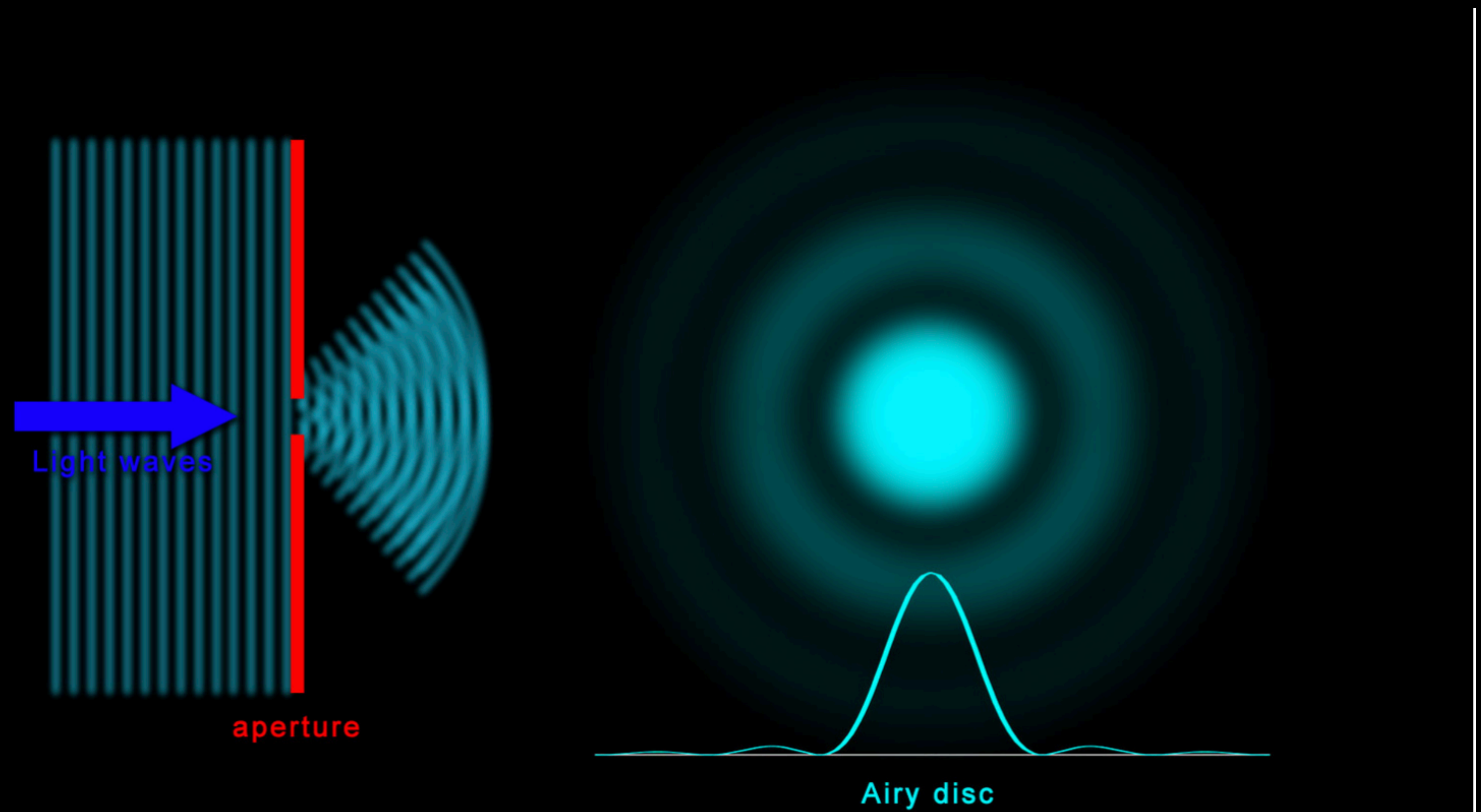
Optics -Design



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:

- Diffraction produces a Bessel function.
- When measured by an image sensor, we obtain the square of the Bessel function, or, the airy disk.
- When we compute the fourier transform of the airy disk, we obtain the optical transfer function (OTF).
- When we compute the magnitude of the (OTF) , we obtain the modulation transfer function (MTF).

Optics - Design

- The MTF quantifies the resolving power of the lens, by plotting line pair contrast versus line pair density / frequency.
- Viewed on its own, the designer would believe that the goal is to maintain line pair contrast while maximizing line pair density / frequency.
- However, there is more to the story.

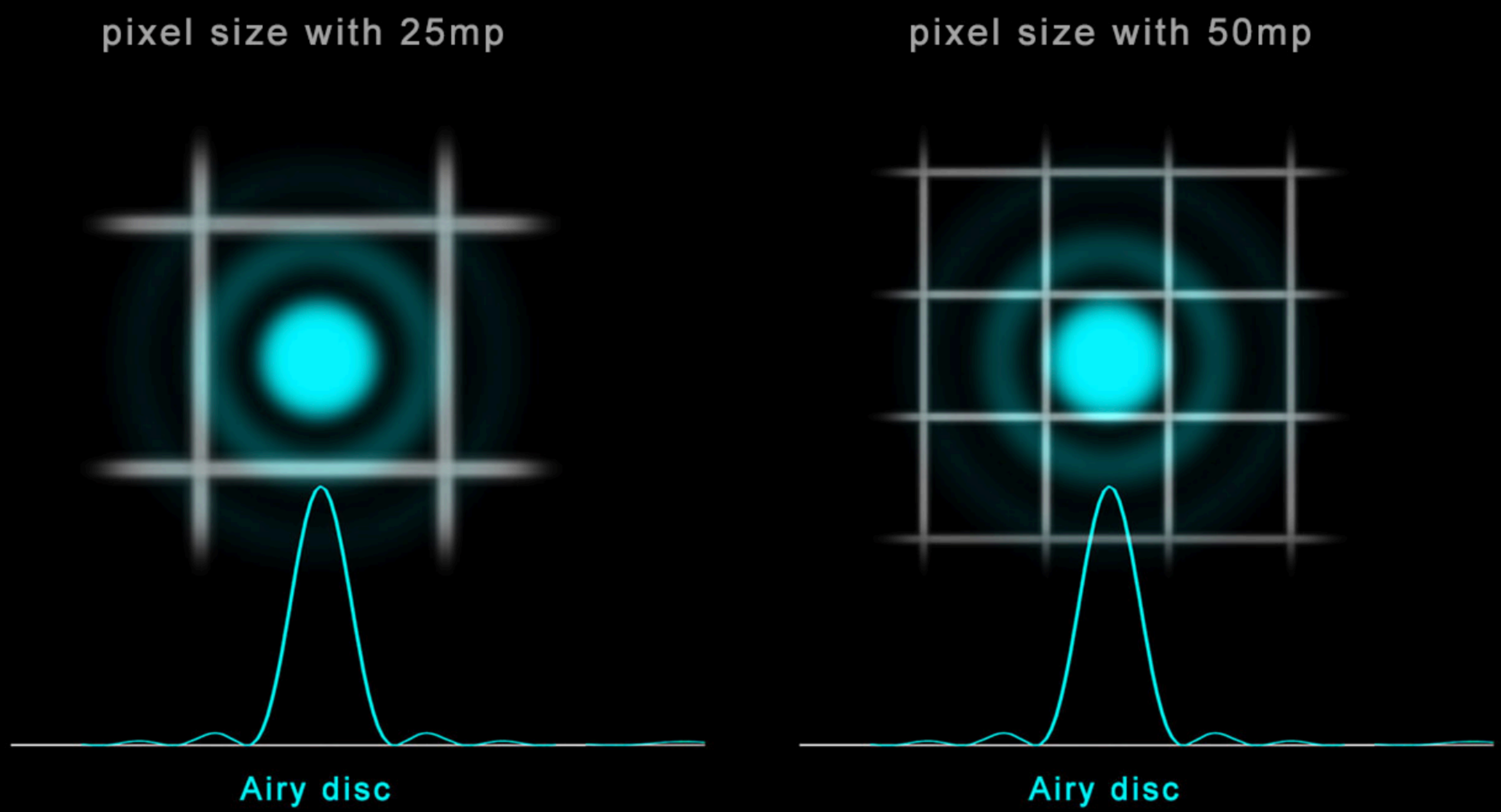
Optics - Design

- The pixel dimension also plays a major role in resolving power.
- Consider a square pixel.
- The square pixel integrates light over a given area.
- Mathematically, the pixel performs a convolution operation with the scene.
- $g(x, y) = I(x, y) * h_{pix}(x, y)$ $\qquad$ $h_{pix}(x, y) = \frac{1}{a * a} rect(\frac{x}{a}) rect(\frac{y}{a})$

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 then given by: $|H_{pix}(f_x, f_y)| = |sinc(af_x) sinc(af_y)|$
- The system MTF is given by: $MTF_{system} = MTF_{diffraction} MTF_{pixel}$
- The picture on the next slide, illustrate the physical dynamics 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

- In terms of the system MTF, we observe that one curve will roll off before the other, setting the line pair density / frequency limit for the system.
- Now, one would think that the best result requires using the largest aperture (to minimize the airy disk diameter) and the smallest pixel dimension (to maximize the pixel MTF).
- Fortunately, this is still not the case.
- But, before an explanation is provided, let’s see what happens, when we try to push the existing limits.

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.
- In addition, smaller pixels mean less light gathering capabilities.
- Less light gathering capabilities means higher noise.

Optics - Design

- Fortunately, there IS one additional factor to consider.
- That factor is sample frequency, and the related concept of aliasing.
- From Nyquist sampling theory, to prevent aliasing artifacts, the sample frequency needs to be twice the frequency bandwidth of the baseband signal.
- Hence, if the bandwidth of the MTF is excessively high and the pixel pitch is excessively wide, aliasing will result - destroying the image quality.

Optics - Design

- To exacerbate matters, sampling on a sensor is also performed using a bayer pattern.
- Hence, the pixel pitch is actually 2x that of a monochrome sensor.
- Given this information, the design procedure should then be:

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 needed, to increase light gathering capabilities.
- With every lens, diffraction will occur, due to a finite lens aperture.
- With every lens, different DoF's will result, based on the focal length of the lens, the focus position, the pixel size, and the aperture size.
- This will result in a wide variety of images, with varying degrees of foreground and background blur.
- Different lenses are designed, to address different use cases.
- The resolving power of the generated image is a function of the lens, the pixel width, and the pixel pitch.