Perceptual Optics

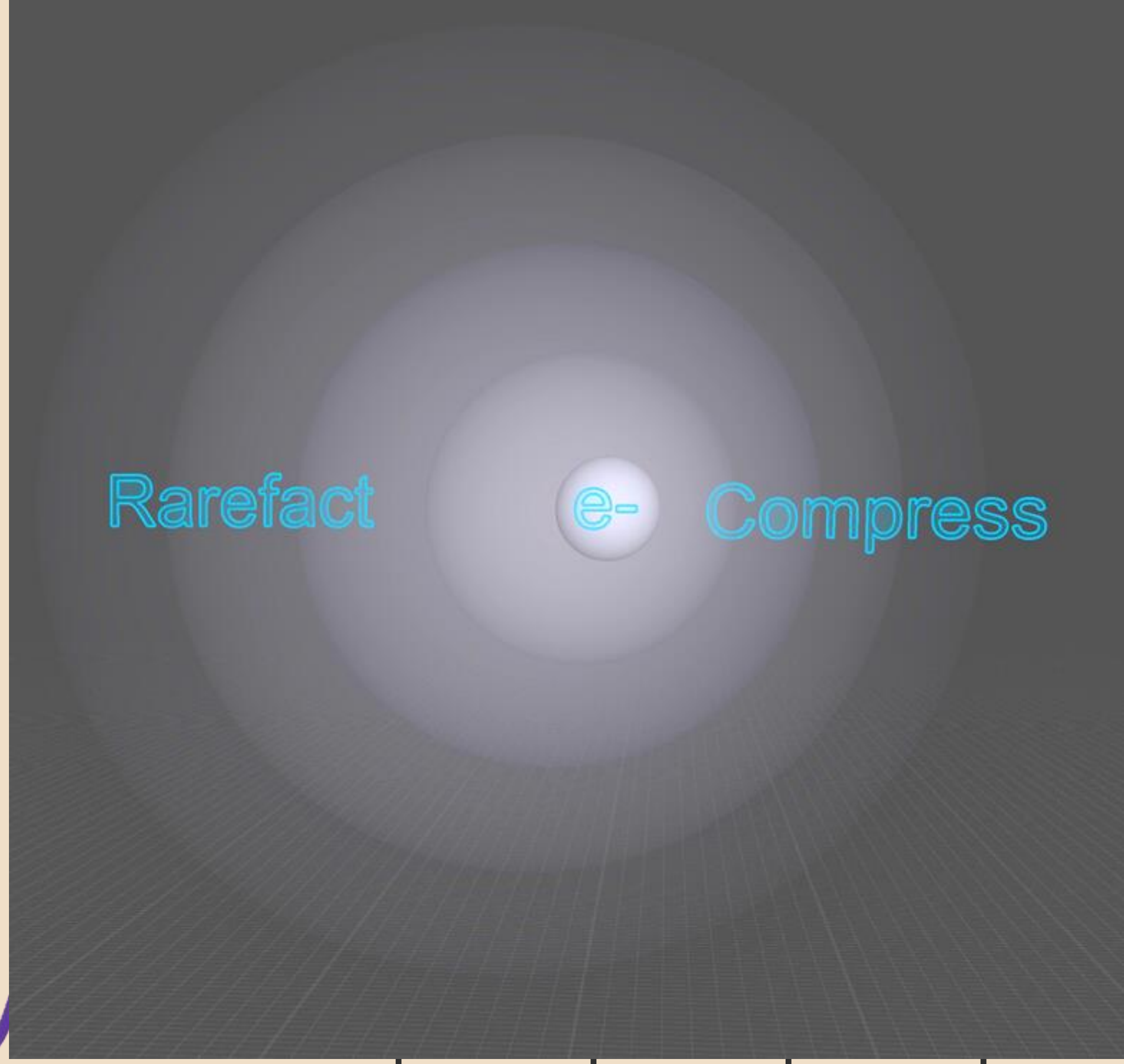
Brad Caldwell

Table of Contents

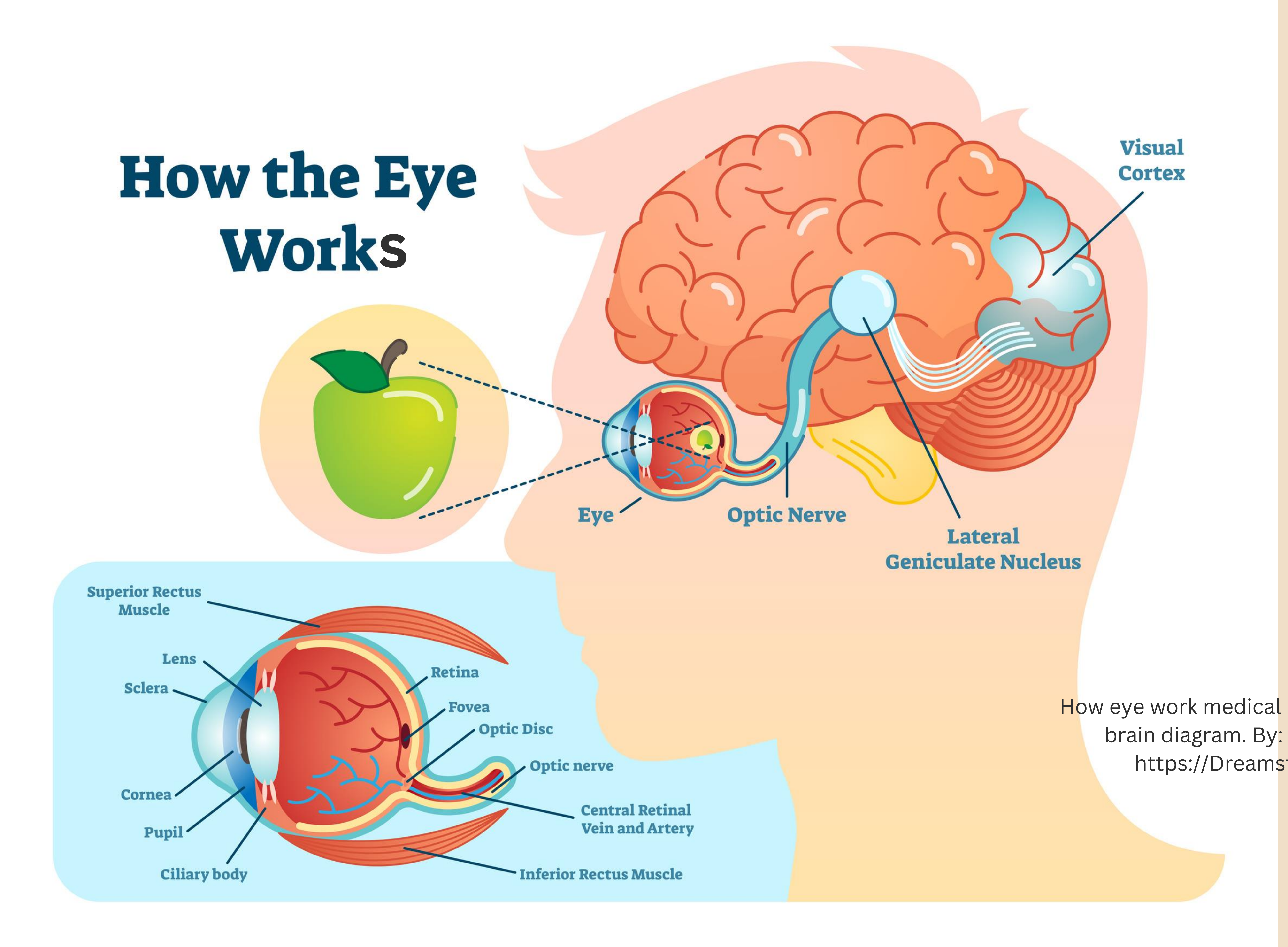
- **Physical and Geometrical Optics**

Light:

From Emission to Capture



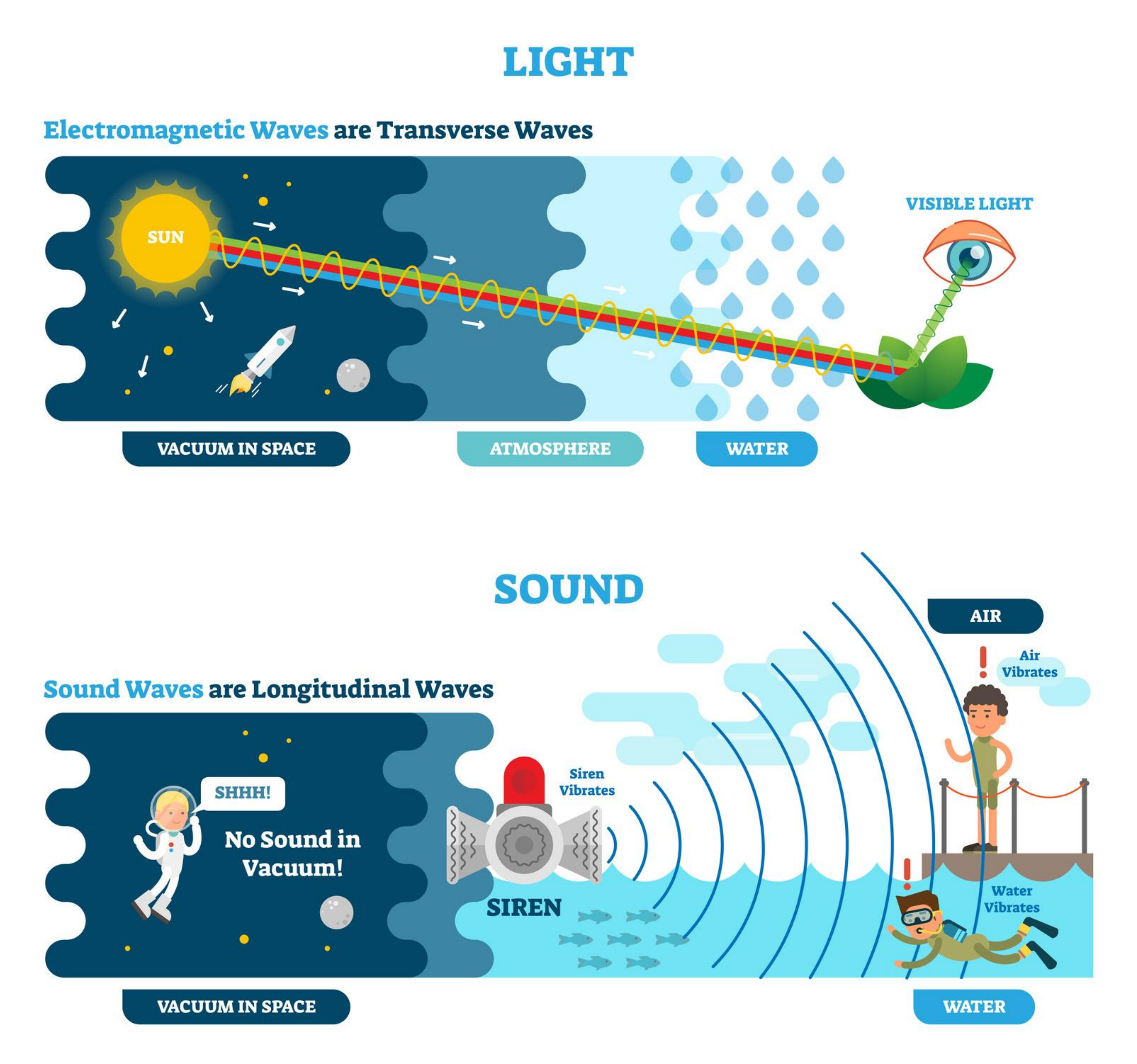
Light is emitted when an electron experiences acceleration. Cyclic vibration in the range of 400-700 THz (400 - 700 trillion vibrations per second) underlie colored light. By compression and rarefaction are meant temporal cyclic perpendicular changes of strength of electric field. Light propagates outward **spherically** at *300 million meters/s*.

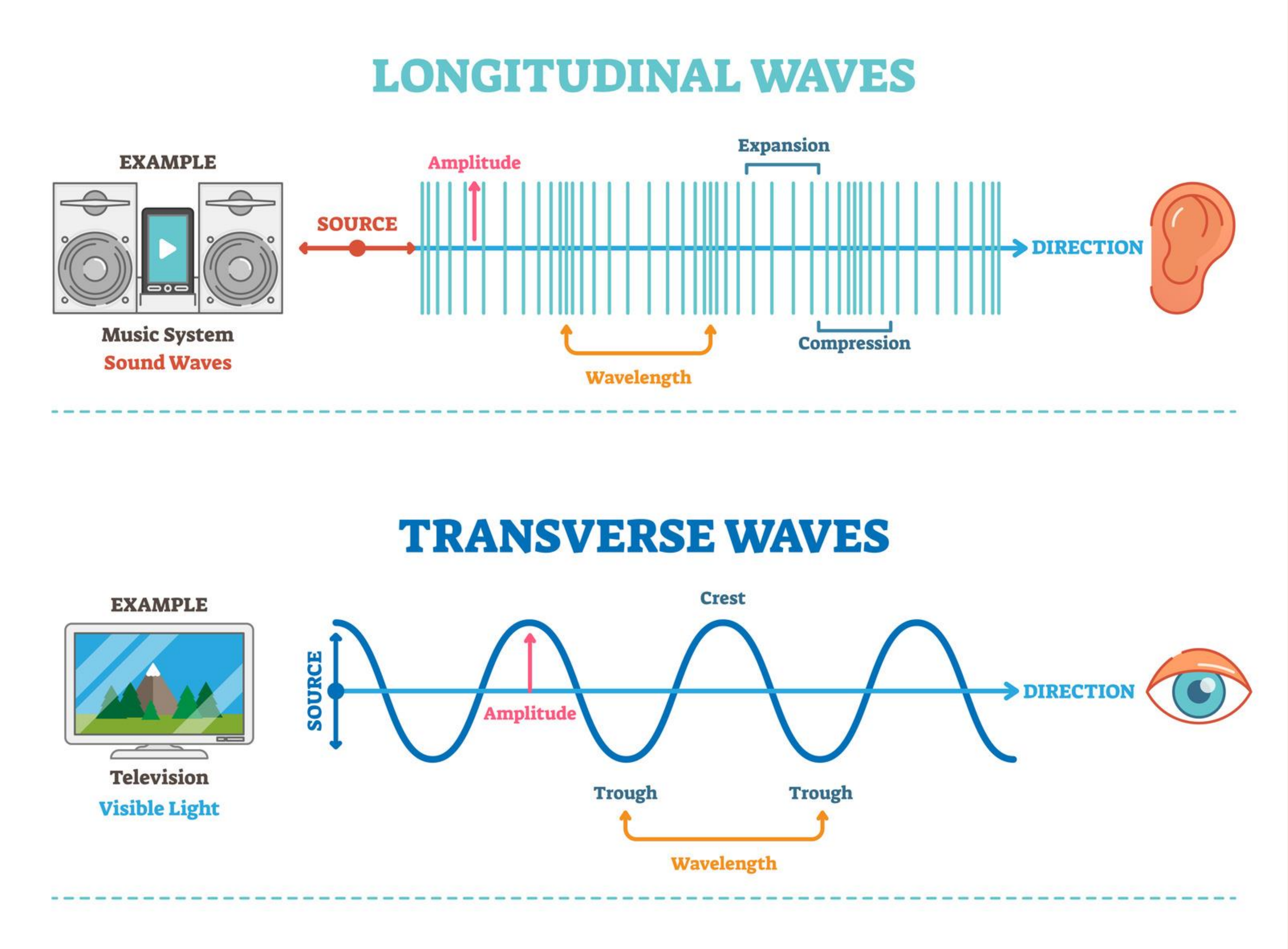


How eye work medical illustration, eye - brain diagram. By: VectorMine. https://Dreamstime.com

Longitudinal vs. Transverse Waves

Longitudinal and Transverse wave type, vector illustration scientific diagram.
Sonic and visual perception principle. By: VectorMine. https://Dreamstime.com





Electro-magnetism

Lienard-Wiechert Equations

$\mathbf{E}(\mathbf{r}, t) = (q / (4\pi\varepsilon_0)) (r / (\mathbf{r} \cdot \mathbf{u})) [(c^2 - v^2)\mathbf{u} + \mathbf{r} \times (\mathbf{u} \times \mathbf{a})]$

$\mathbf{B}(\mathbf{r}, t) = (1/c) (\mathbf{r} \times \mathbf{E}(\mathbf{r}, t))$

$\mathbf{E}(\mathbf{r}, t)$ is the electric field at position r and time t
q is the charge of the source
ε_0 is the vacuum permittivity
r is the vector from the retarded position of the source charge to the impacted charge; r is scalar
u is the unit vector along **r** (**u** = **r**/r)
v is the speed of the source charge
a is the acceleration vector of the source charge
c is the speed of light

Griffith's Introduction to Electrodynamics

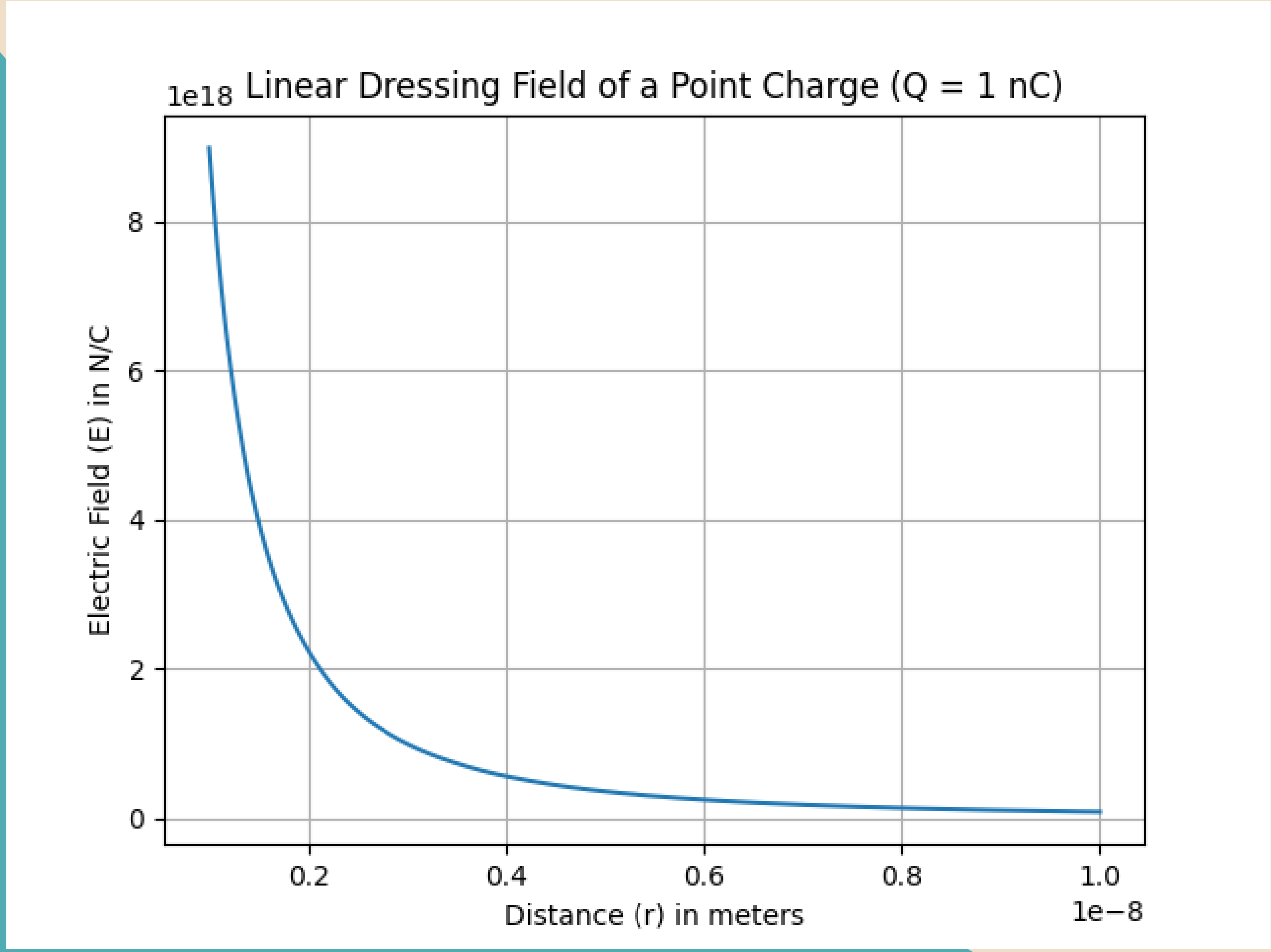
In classical electrodynamics, a stationary or moving charge has a "dressing field." An accelerating charge emits a "radiation field" at the speed of light.

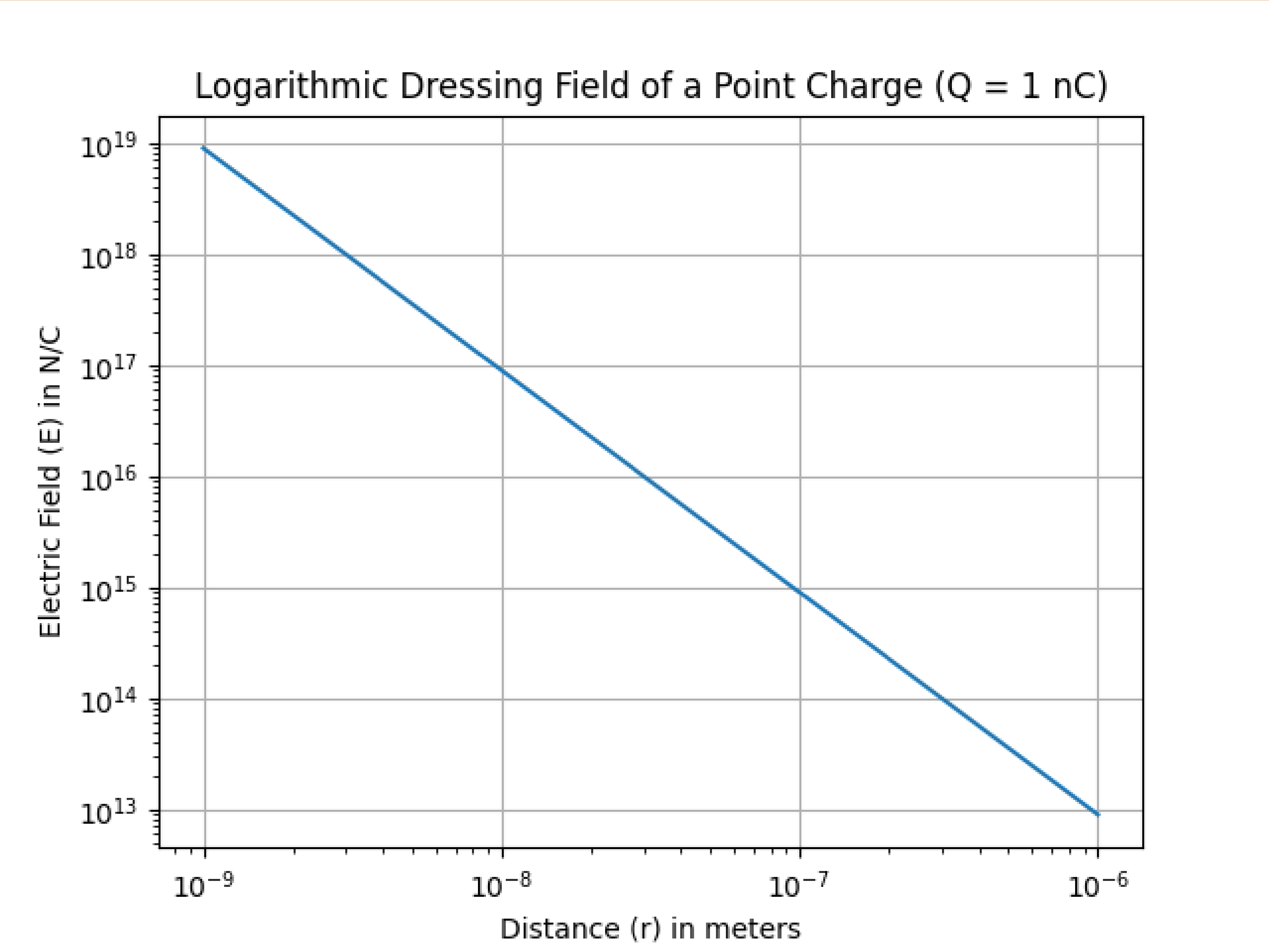
- When the charge is *stationary* or *moving at a constant velocity*, the Lienard-Wiechert potentials describe the (electric & magnetic) **dressing field**.

- When the charge is *accelerating*, the Lienard-Wiechert potentials account for the (electric & magnetic) **radiating field**, which propagates as an electromagnetic wave.

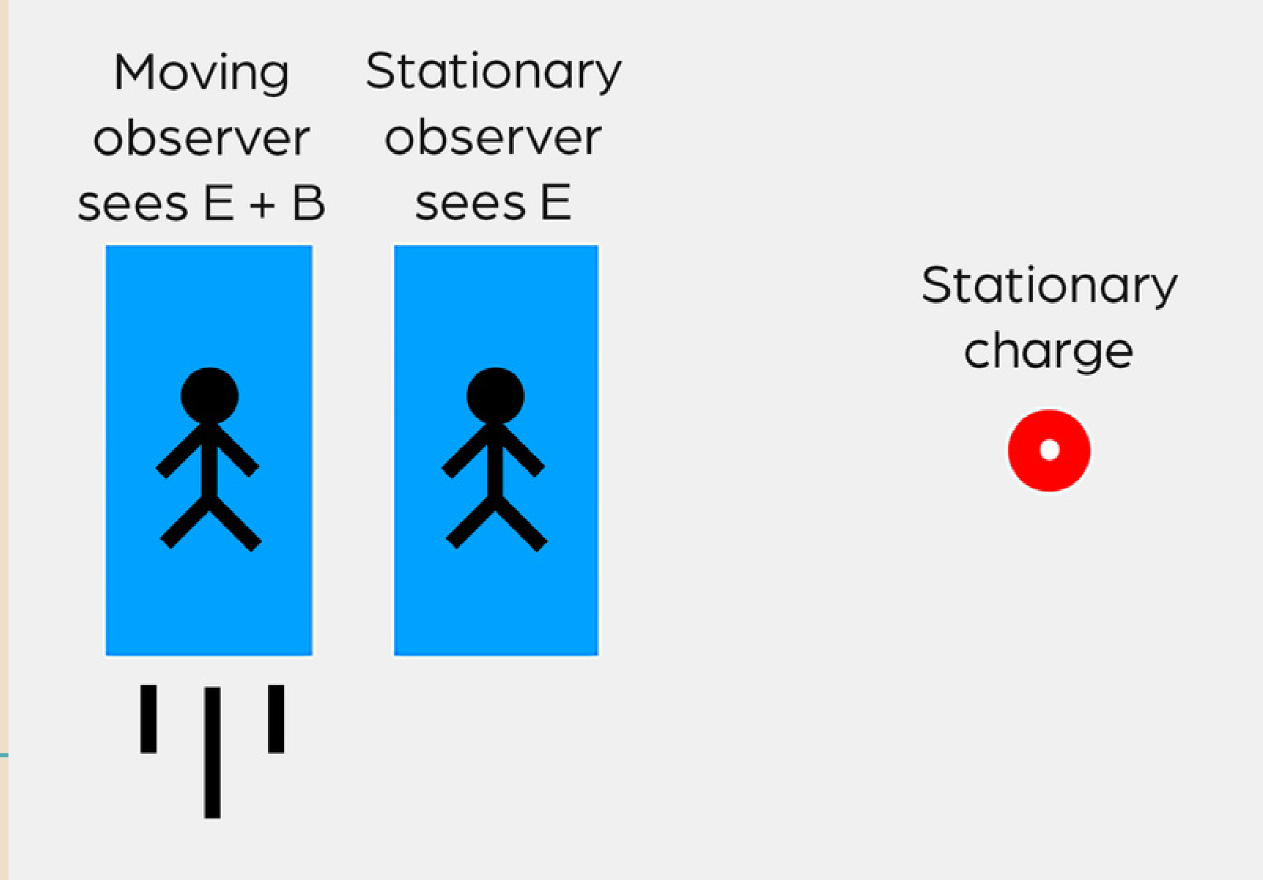
A Charge's Dressing Field

When a charge is not accelerating, the field around it "hangs around." Even when the charge is moving at a steady rate, it still just moves with it. Only when it accelerates does it cause a radiating field (electromagnetic waves; light). From Coulomb's law: **$E = k * Q / r^2$** is the static **dressing field**.





Einstein's Special Relativity



(*Accelerating* observer would experience Unruh effect [bath of particles])

One of the key results of Einstein's special relativity is the **Lorentz transformation**, which describes how the electric and magnetic fields transform between different inertial frames. Special relativity shows that an observer in one inertial frame may see a purely electric field, while another observer in a different inertial frame moving with respect to the first observer may see a combination of electric and magnetic fields. *This further highlights the interdependence of electric and magnetic fields and their common origin.*

Quantum Electrodynamics

Does quantum consider electric and magnetic similar?

Dressing field *= "virtual photons"*
Radiating field *= "real photons"*

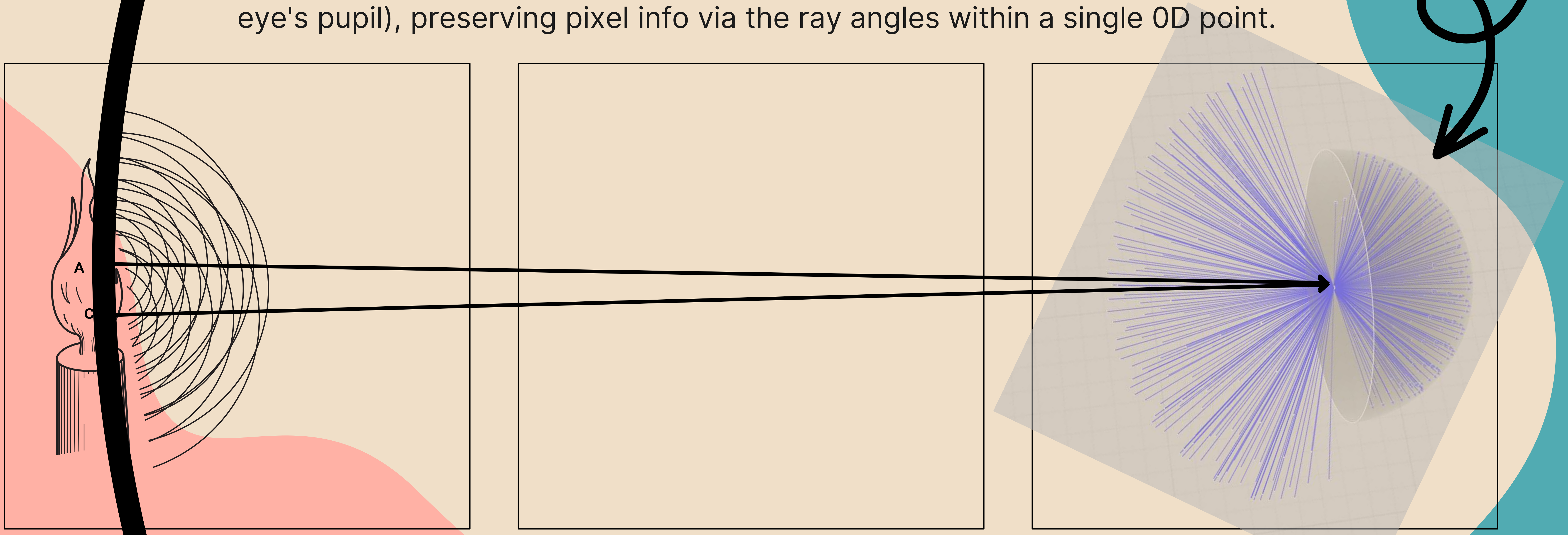
In the context of quantum mechanics, electromagnetic fields and waves are described by the quantization of the electromagnetic field, which leads to the concept of **photons**. Photons are the elementary particles that represent quanta of the electromagnetic field, including electromagnetic radiation. Quantum electrodynamics (QED) is the theory that describes the interaction between electromagnetic fields and charged particles (such as electrons) in terms of the exchange of photons. *While QED is fundamentally different from classical electrodynamics, it still involves the interplay between electric and magnetic fields and their common origin in charged particles.*

Spheres

Weird pinhole camera

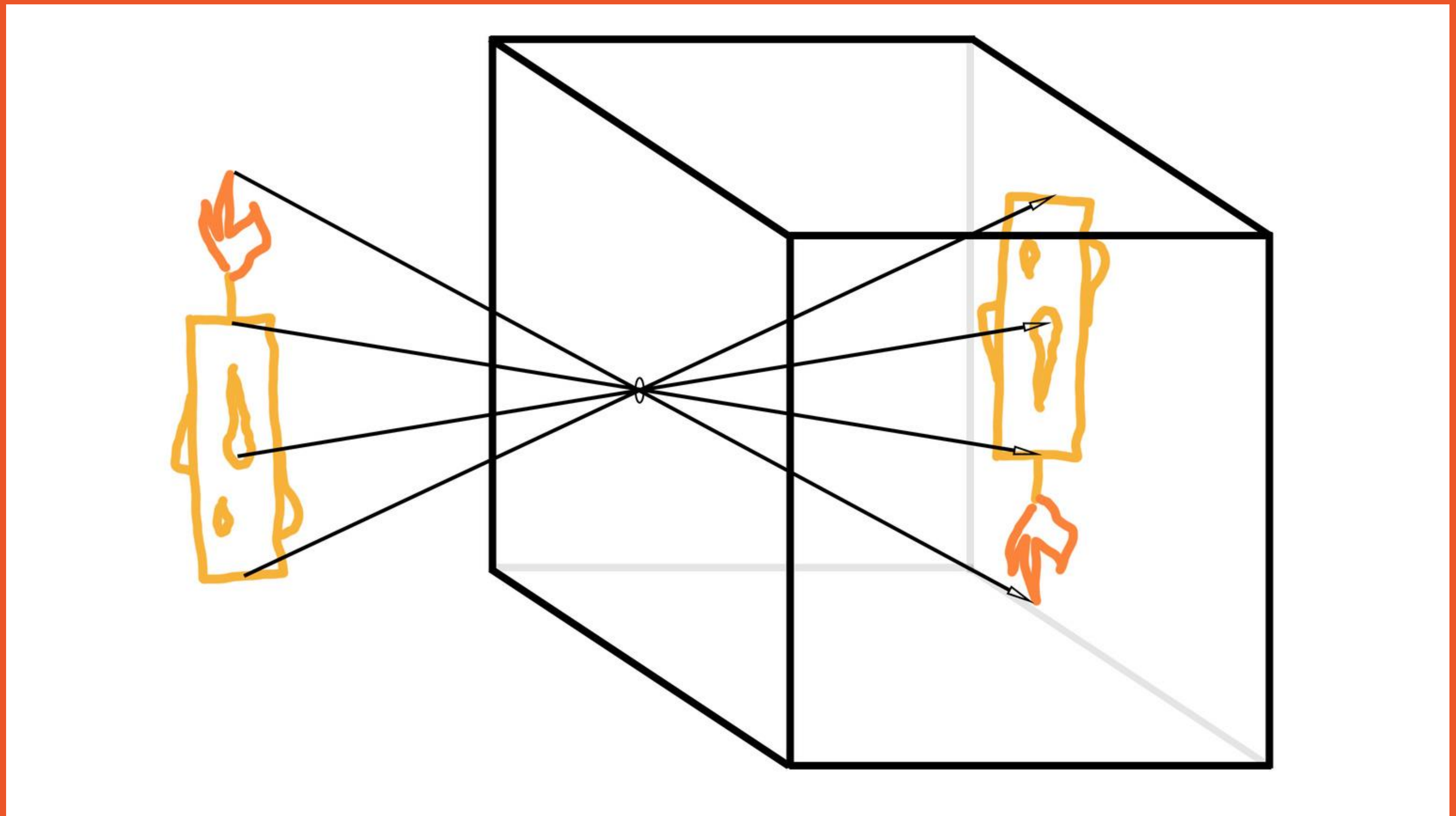
Light emits spherically from each of a multitude of points.

Taking an image eliminates radial distance info from each point in space by collapsing focal sphere shells onto a camera's pinhole or lens (or the eye's pupil), preserving pixel info via the ray angles within a single 0D point.



Emission

Capture



Pinhole Cameras

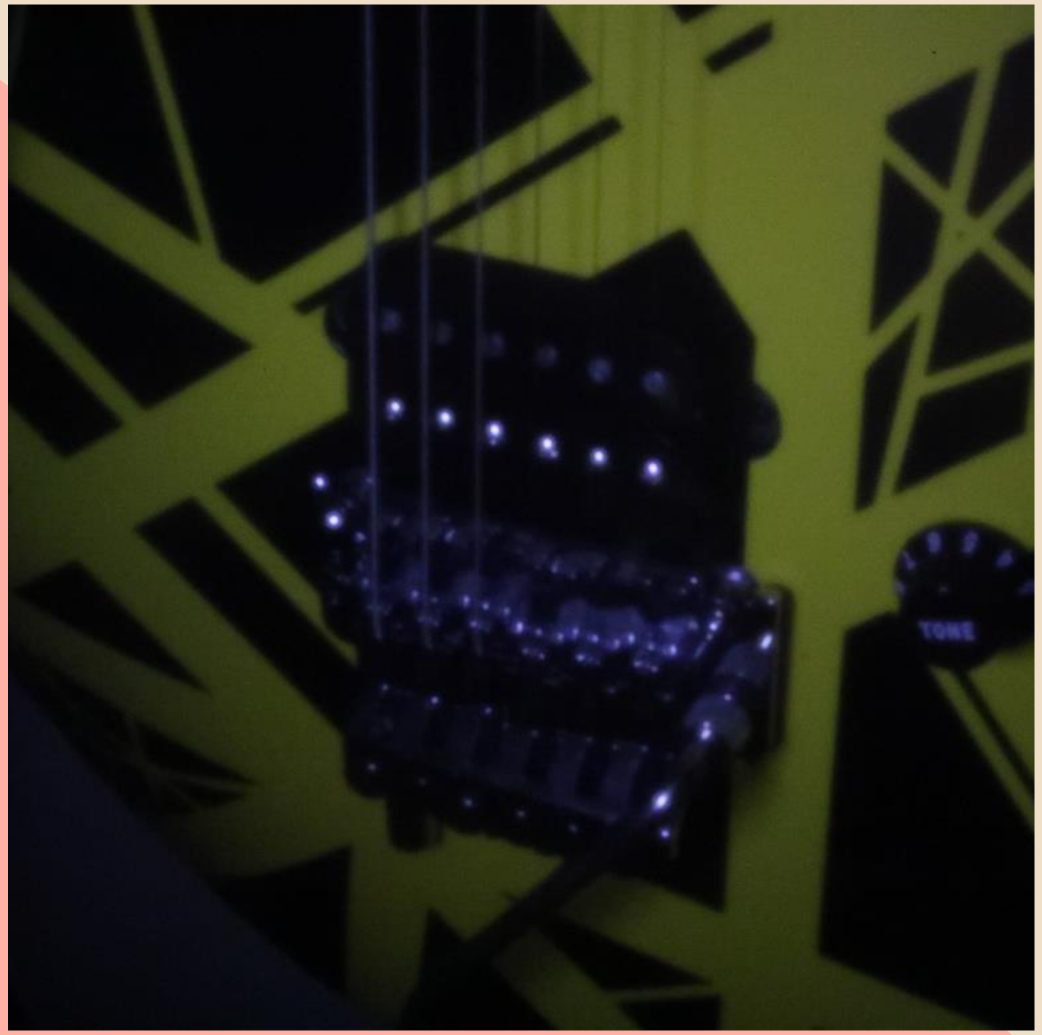
When examining the physics of photography, one must first conceptualize the existence of spatial points within the environment and a pinhole aperture. Establish a connection between each point in space and the pinhole, and subsequently extend these lines to reach the camera's sensor located at the back of the device.

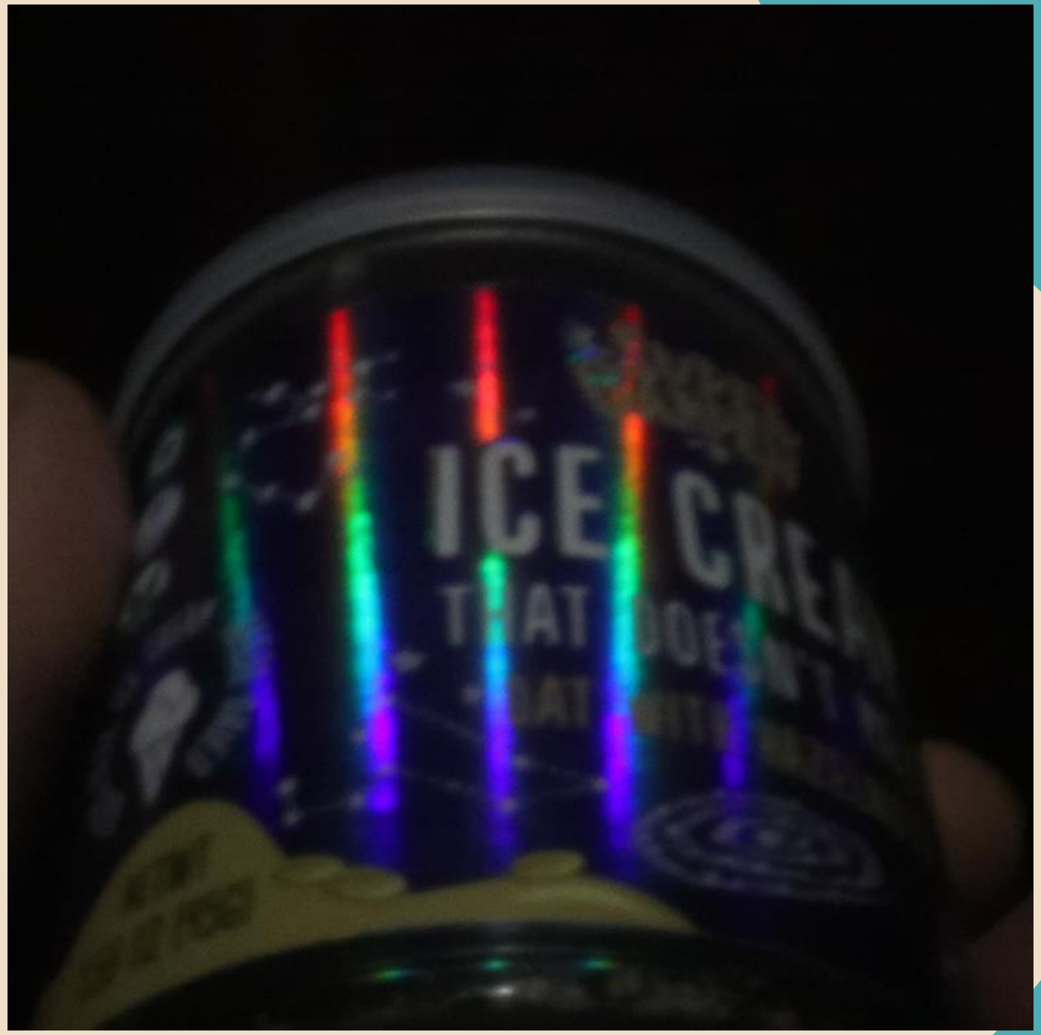
https://www.quora.com/What-is-a-pinhole-camera

Pinhole Cameras

pinhole images

The benefit of pinhole "lenses" or openings is that there is no need to focus. However, a little blurriness in inevitable (aperture still too large or will cause diffraction when made very small), and a lot of light is required.

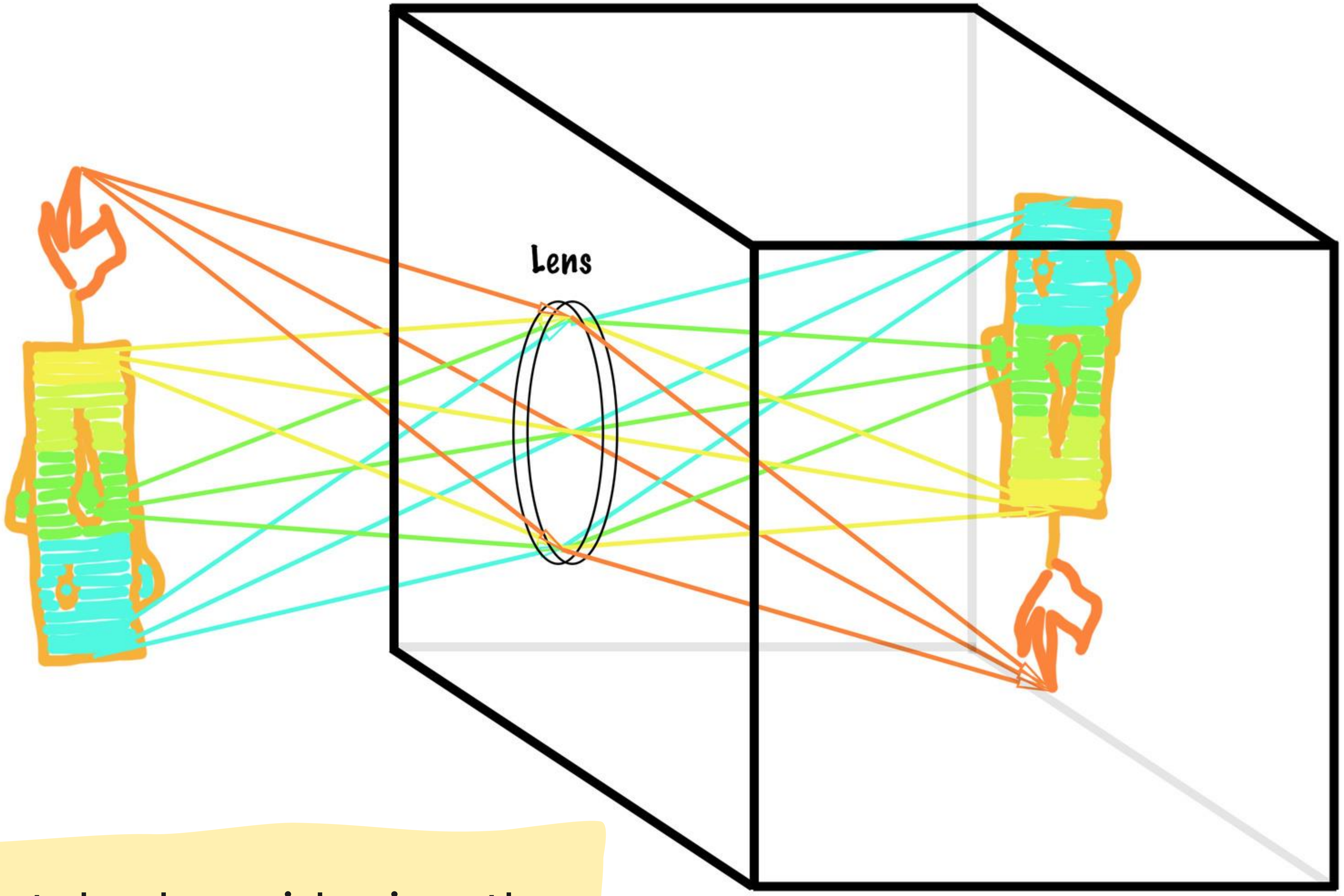




Pinhole Summary

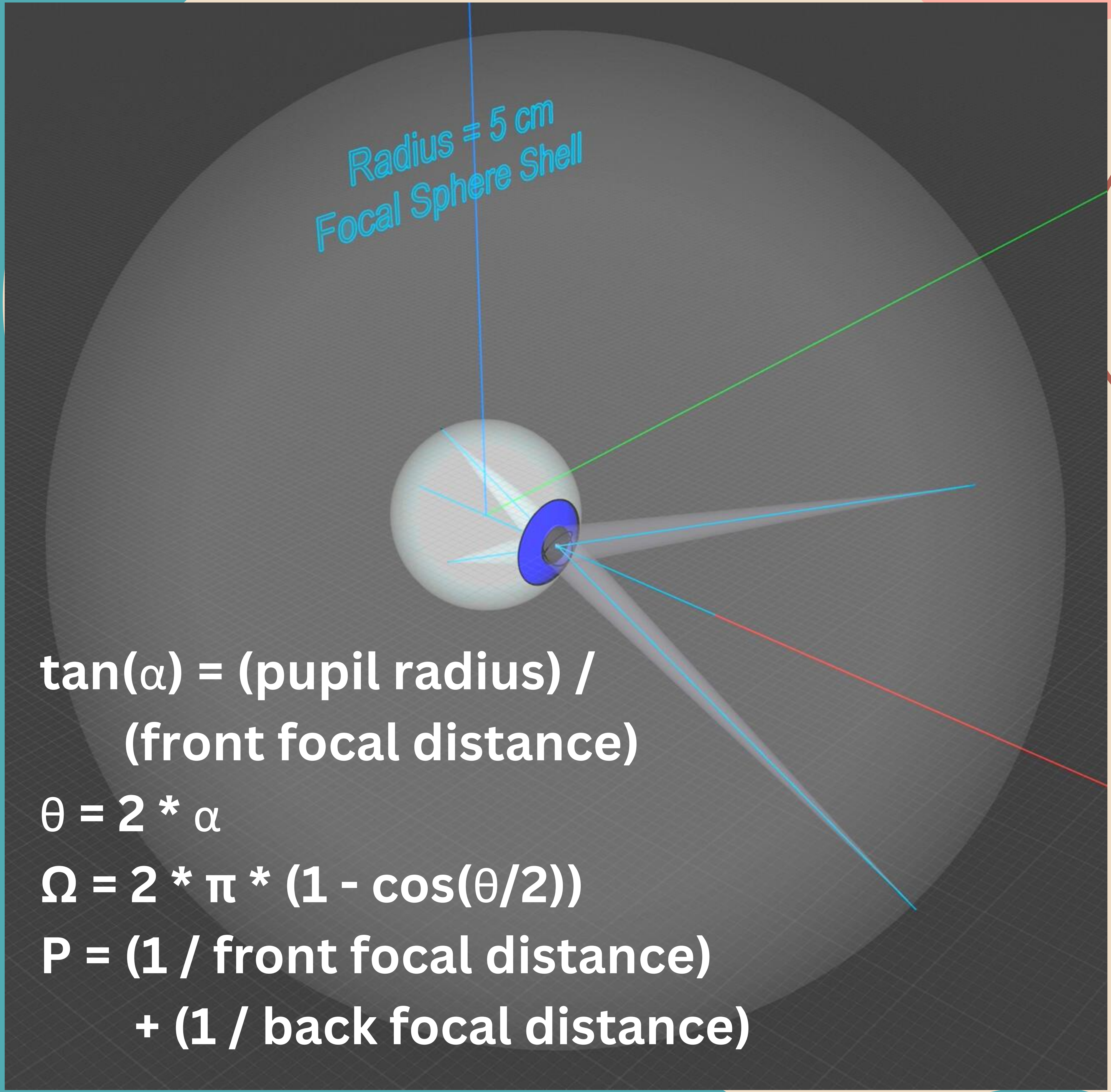
- Pro: No need for focusing mechanisms, simplifying the design.
- Con 1: Requires a substantial amount of light, limiting low-light performance.
- Con 2: Cannot achieve perfect clarity due to diffraction effects at the very small apertures that would otherwise produce clarity (~0.1 mm); optimal aperture size is approximately 0.35 mm.

Lenses



Lenses enable increased light intake by widening the aperture. As they admit a conical spread of light from each point in space, they must **focus,** or *refract*, the light rays to converge at a single point on the sensor.

Front Focal Distance



$\tan(\alpha)$ = (pupil radius) / (front focal distance)

$\theta = 2 * \alpha$

$\Omega = 2 * \pi * (1 - \cos(\theta/2))$

P = (1 / front focal distance) + (1 / back focal distance)

A front focal distance of **5 cm** (5 degree, i.e., *27 square degree* solid angle cone) would require **70 diopters** of *focal power* (assuming 4 mm pupil diameter and 2 cm from lens to retina) - too close to focus on for the eye's range (50-60 diopters).

Front Focal Distance

In contrast, a **3 cm** focal distance means:

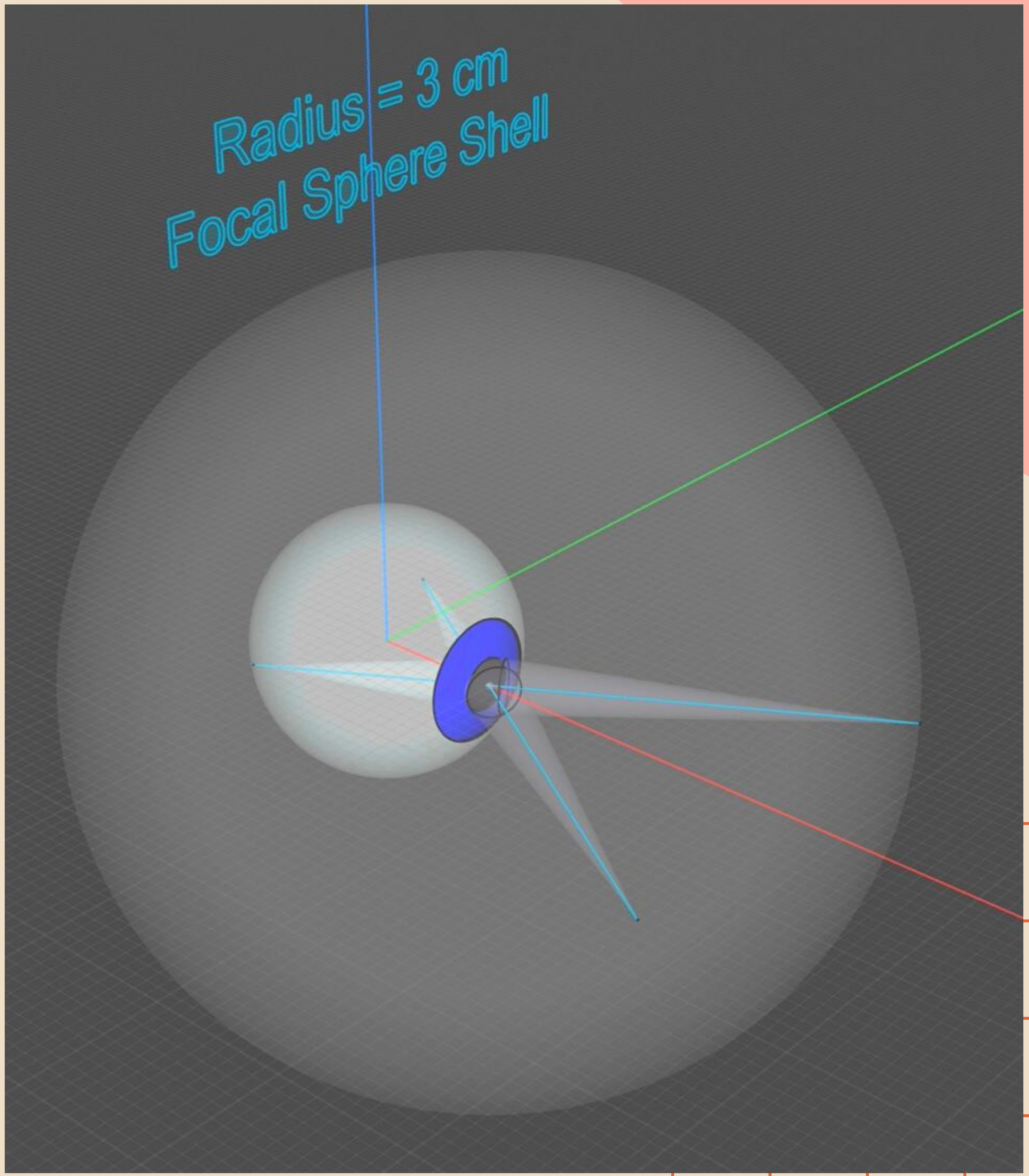
- θ ≈ 7.5 degrees; Ω ≈ *100 square degrees*
- P ≈ **83 diopters** (refracting focal power)

Similarly, a **1 cm** focal distance would mean:

- θ ≈ 23 degrees; Ω ≈ *935 square degrees*
- P ≈ **150 diopters** (refracting focal power)

Also, a **100 cm** focal distance would mean:

- θ ≈ 0.2 degrees; Ω ≈ *0.004 square degrees*
- P ≈ **51 diopters** (refracting focal power)



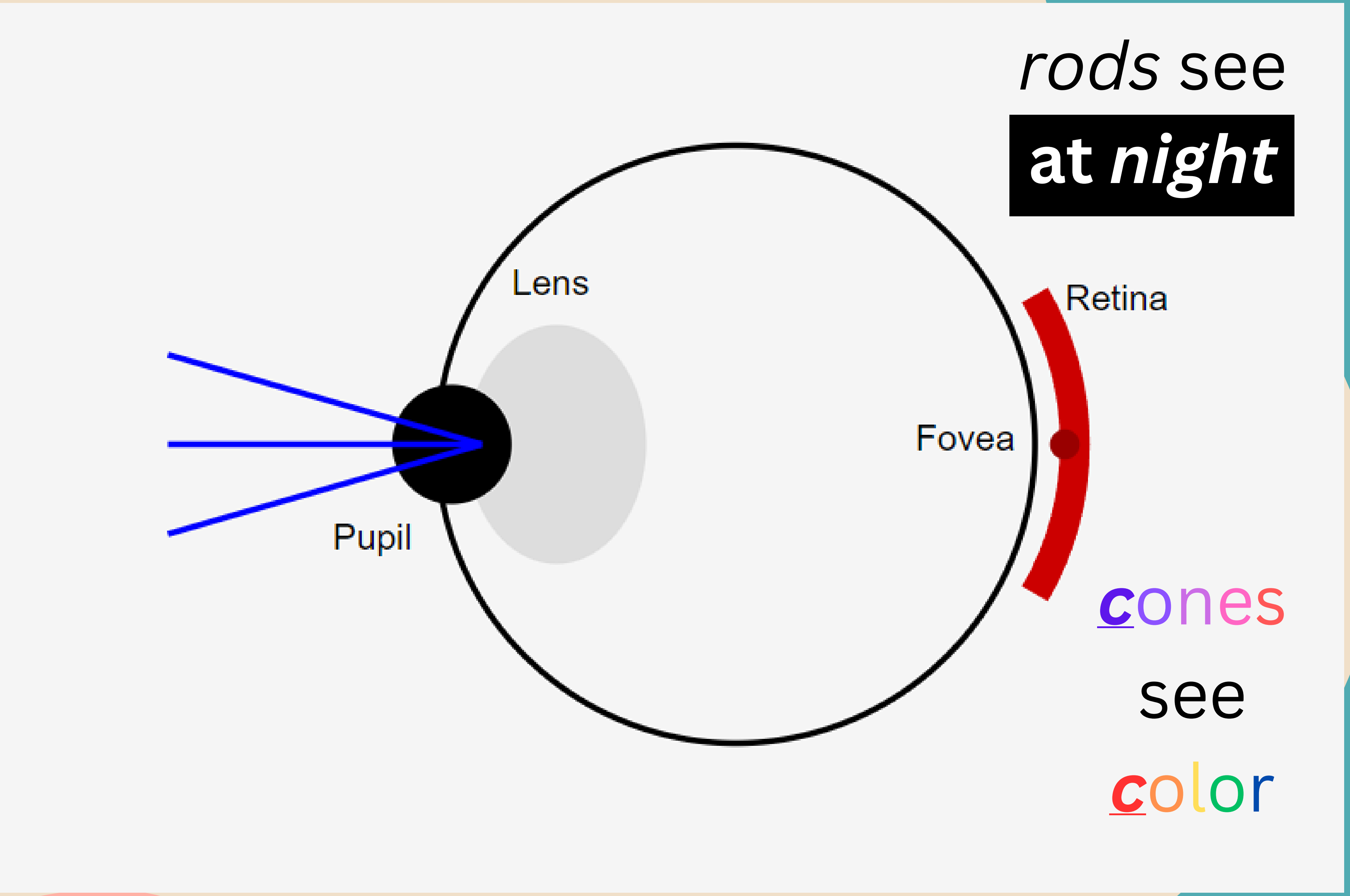
Near/Farsightedness

In **nearsighted** individuals (**myopia**), the eye has *too much focusing power*. This causes light entering the eye to be focused in front of the retina rather than directly on it, leading to blurry vision for distant objects.

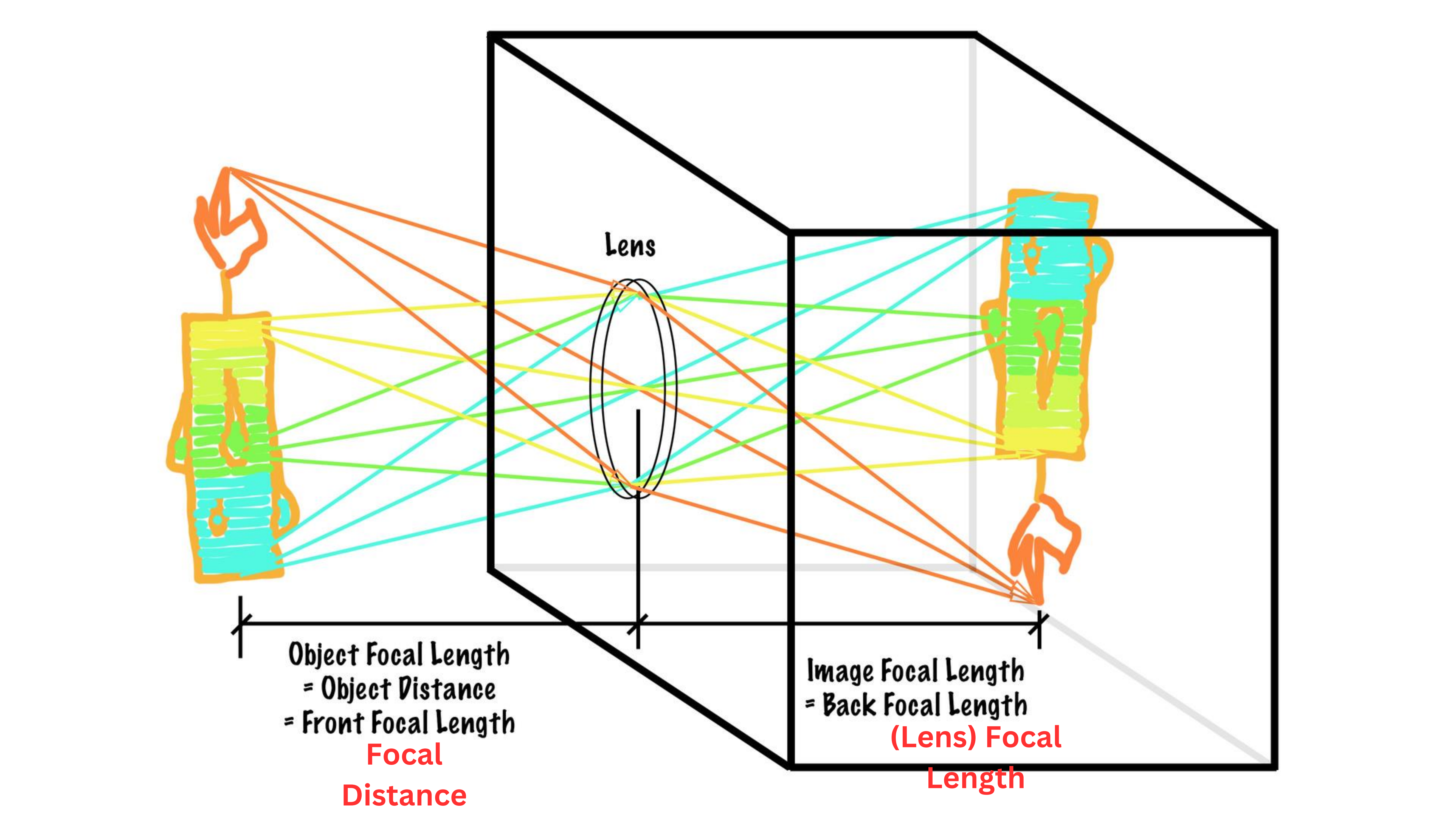
With the cornea providing about 40 diopters and the lens, in tension, providing 10 diopters; we get 50 total for focus on far objects/infinity. When lens relaxes, you get 55-60 diopters total, allowing focus on nearer objects. By the age of 40-45 years, the average accommodation reduces somewhat. This reduction in accommodation ability is known as **farsightedness** (**presbyopia**) and results in difficulty focusing (*insufficient focusing power*) for near objects.

The Eye

The cornea at the front of the eye provides ~40 diopters of focusing power. The adjustable lens provides adjustable range (10-20 diopters for young adults, which may decline to 10-15 diopters adjustable range at older age).



Focal Vocabulary



Focus Breathing

Although a lens has a set **focal power** which is expressed, for convenience, as the distance between lens and sensor, when you focus on items shorter than infinity, you are slightly increasing the distance between the lens and the sensor, allowing more distance for the pixels from closer items to converge into a point before reaching the sensor. This blurs the background, but also *decreases the field of view* (*increases zoom*). This is known as **focus breathing**, and it is demonstrated here for a 100mm lens on a full frame camera, considering the diagonal field of view in degrees.

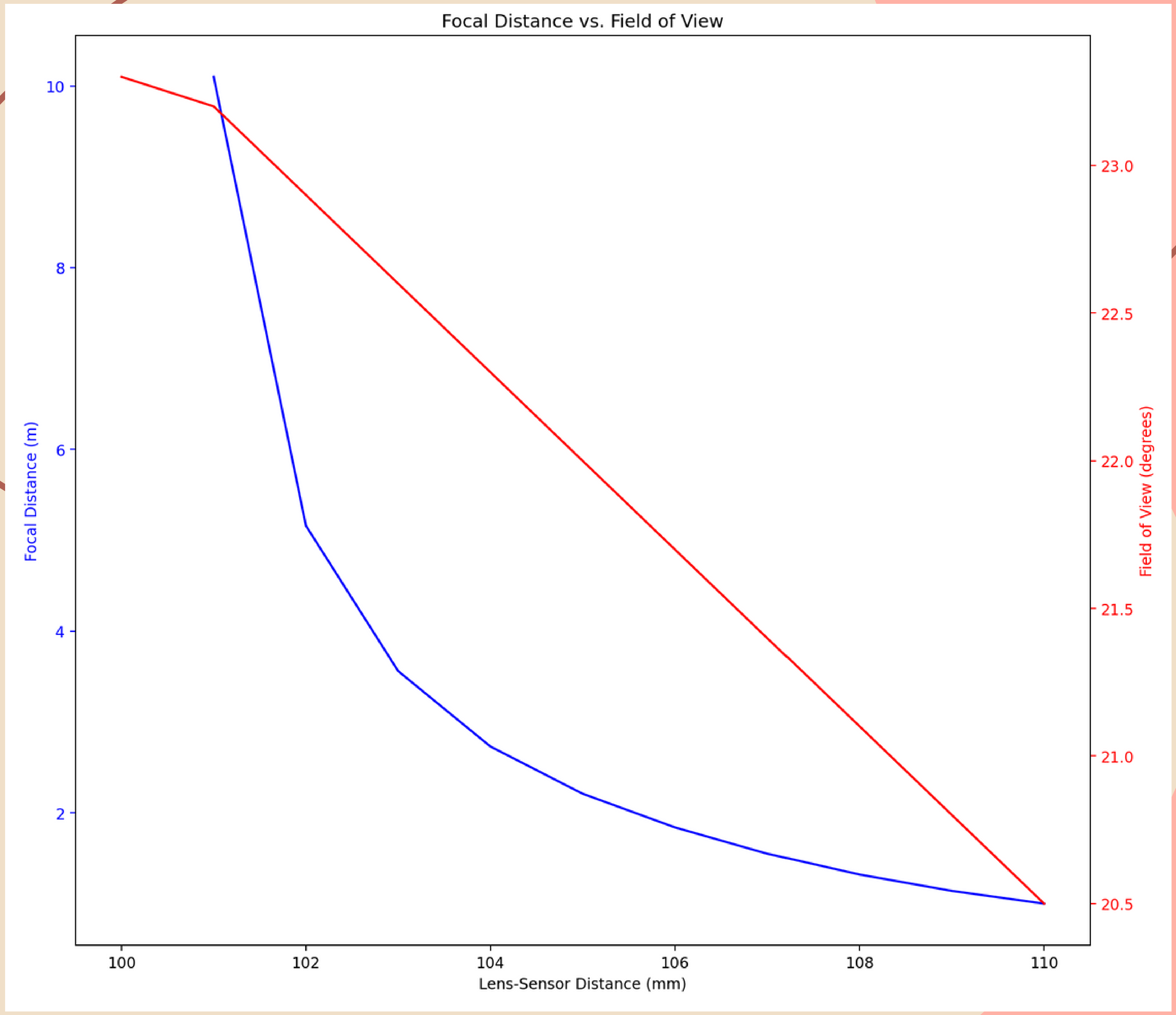
Cine lenses have a way to essentially eliminate focus breathing.

Focal Distance vs. Field of View

Subtitle: Assumes full frame camera, 100mm lens

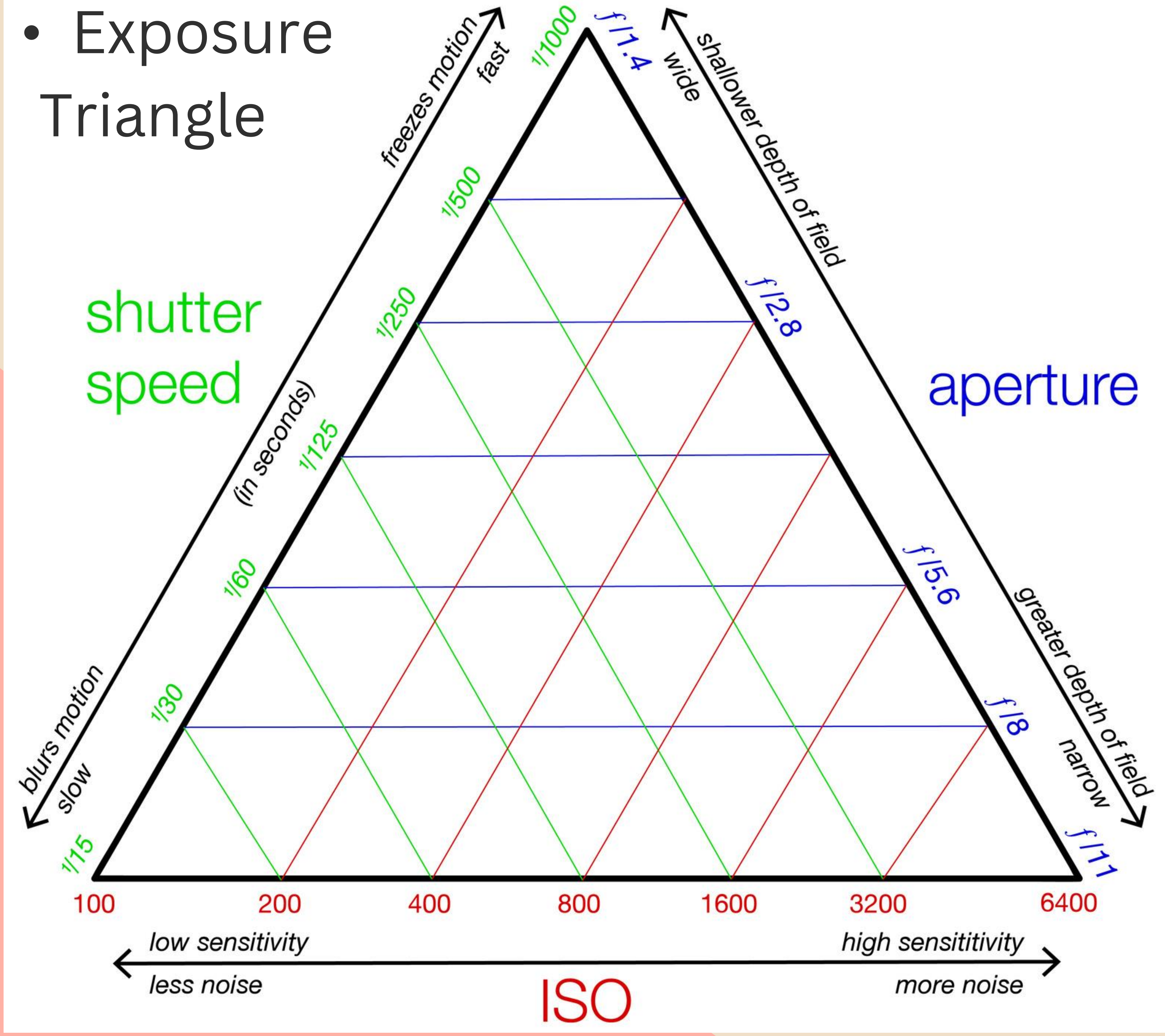
Lens-Sensor Distance (mm)	Focal Distance (m)	Field of View (degrees)
100	Infinity	23.3
101	10.1	23.2
102	5.16	22.9
103	3.56	22.6
104	2.73	22.3
105	2.21	22.0
106	1.84	21.7
107	1.55	21.4
108	1.32	21.1
109	1.14	20.8
110	1.00	20.5

Please note that these values are approximate, and the actual results may vary depending on the specific camera and lens used.



Photography Variables I

- Exposure Triangle



ISO, **aperture**, & **shutter speed** vary exposure.

Depth of field refers to the range of distance within an image that appears acceptably sharp. A shallow depth of field results in a small area of the image being in focus, while a deep depth of field keeps more of the image in focus. Depth of field is affected by factors such as aperture, focal length, and the distance from the subject.

Metering modes determine how the camera measures light to calculate the correct exposure settings. Different metering modes consider different parts of the image, such as the center, the entire frame, or specific focus points.

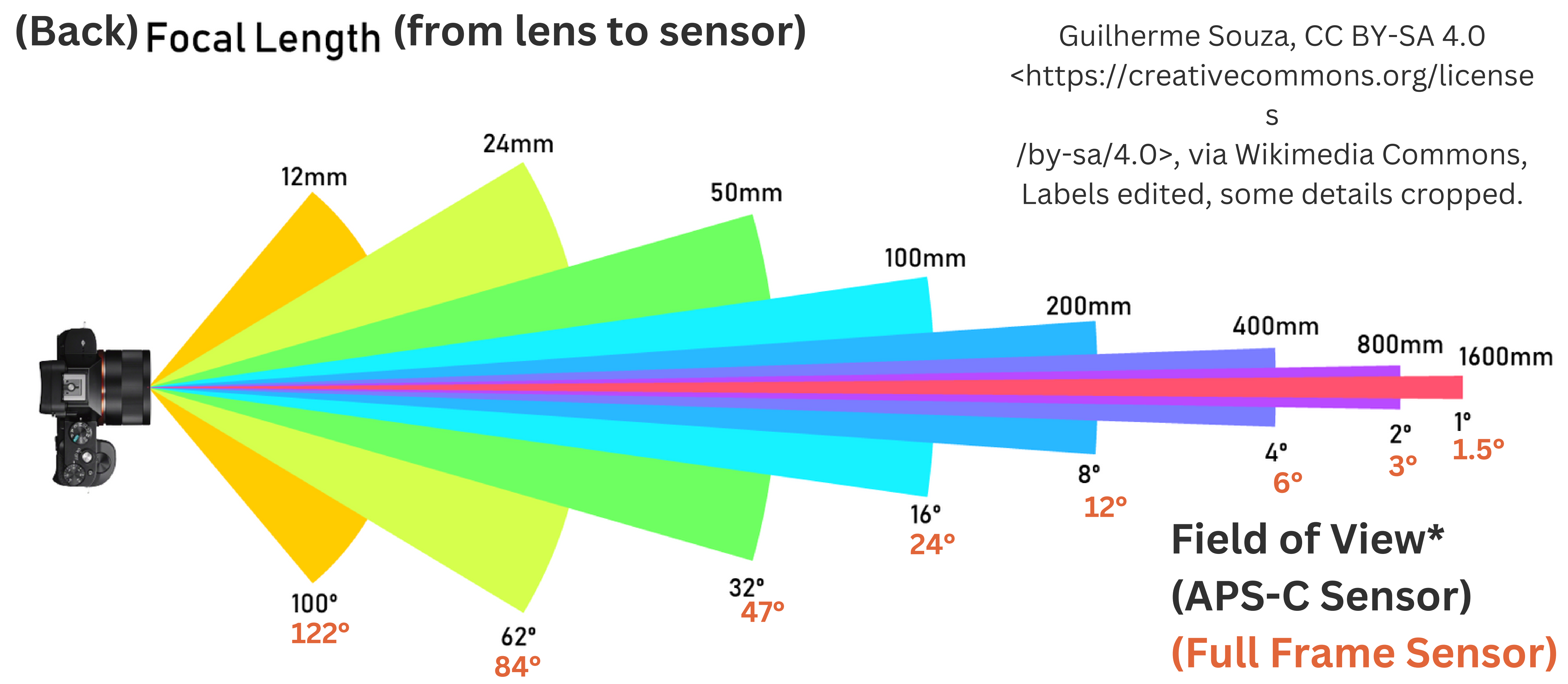
Photography Variables II

- The **focus distance** setting determines which part of the image appears sharp. Autofocus systems use various methods to determine the correct focus, while manual focus allows the photographer to choose the focus point themselves. Many cameras allow for moving default auto-focus point off center when desired.
- **Field of view** is inversely related to *(back) focal length* (from lens to sensor).

$\theta = 2 * \arctan(d/2f)$

d is diagonal along sensor (or horizontal, or vertical),
f is (back) focal length (from lens to sensor).

*In the image, the diagonal (as opposed to horizontal or vertical) distance along sensor was used (so, diagonal angle field of view shown).

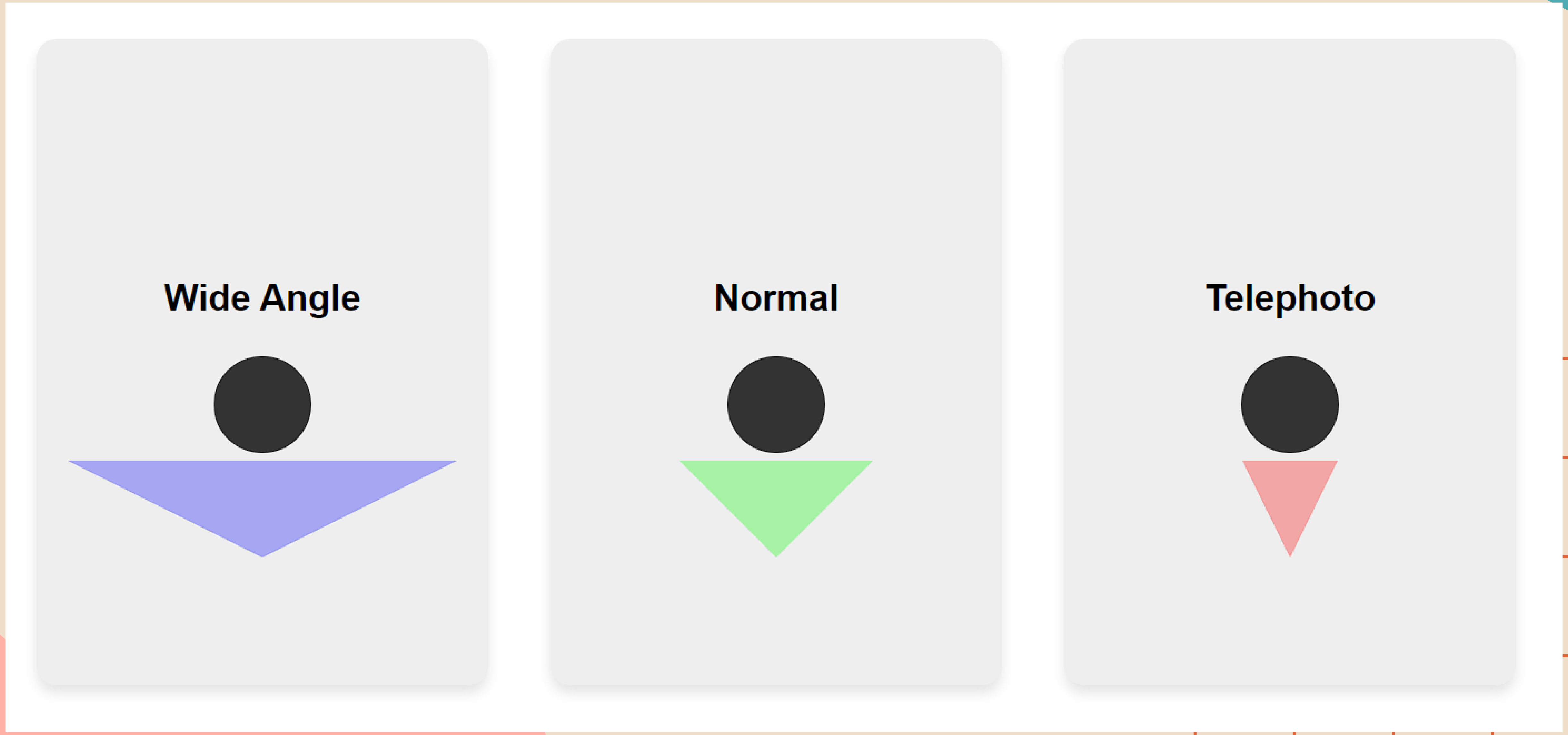


Photography Variables III

tell me about field of view (back focal lengths)

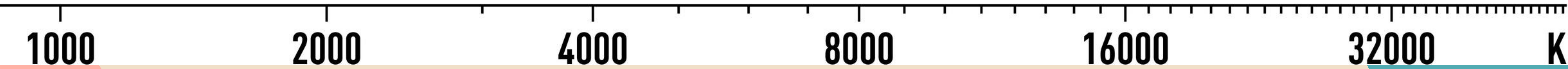
Three broad categories of different (back) **focal length** lenses, with their various *fields of view*, are **wide angle** (< *35 mm*), **normal** (*35 mm - 70 mm*), and **telephoto** (> *70 mm*).

Website table listing fields of view for (back) focal lengths for various sensors:



Photography Variables IV

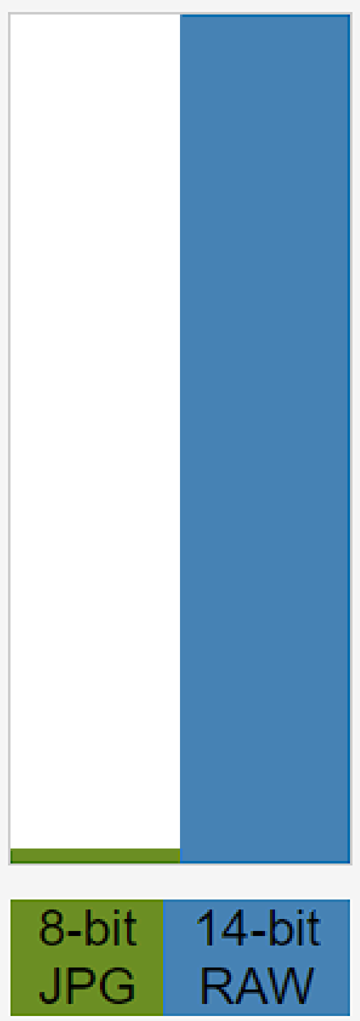
White balance (top), color depth (bottom right)



- **White balance** is the adjustment of color temperature in an image to produce accurate colors. Different light sources have different color temperatures, which can affect the appearance of white objects in the photograph. By setting the correct white balance, you can ensure that white objects appear white and other colors are rendered accurately. **Color temperatures** *in Kelvin* shown across header image; you can adjust this *to approximate lighting source* as one way to set white balance.

- Different cameras offer different **image modes**. For example, JPEG is a common compressed file format, while RAW files are unprocessed and contain more image data. The Canon 5D Mark IV supports 14-bit color depth in RAW mode, which provides more color information and dynamic range than a JPEG image. A 14-bit image would need to use 14-bit depth in Photoshop to preserve details.

- **Color depth** refers to the number of bits used to represent the color of a single pixel in a digital image. A higher color depth allows for more distinct colors and smoother gradients between colors. 14-bit color depth can represent 16,384 intensity increments for each color channel (red, green, and blue), compared to 8-bit's 256 increments (see bar graph).

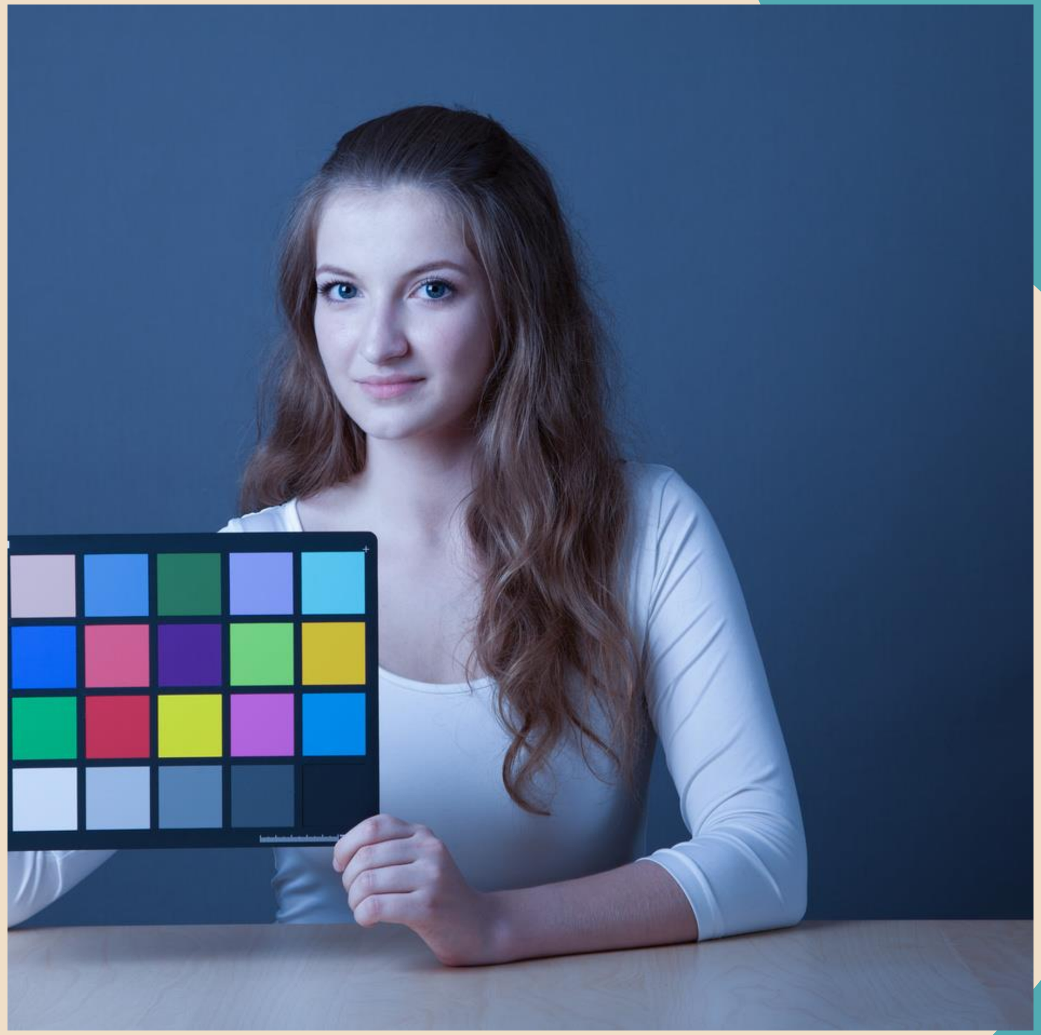


White Balance

Illustrated here are the effects of modifying white balance settings in relation to constant light temperature conditions. If the photographer had utilized warm, neutral, and tungsten lighting, and adjusted the camera's white balance to correspond with each, the resulting images would have appeared virtually identical. The noticeable differences between the photos demonstrate that a single, unchanging light source was employed, while the white balance settings were adjusted. This manipulation of white balance settings significantly transformed the appearance of the images from one configuration to another.

Warm

Neutral



Tungsten

Videography Variables

- **Frame rate**: Videography requires the recording of multiple frames per second (fps) to create the illusion of motion. Frame rates vary depending on the desired look and feel of the video, with common frame rates being 24, 30, and 60 fps. In contrast, photography captures a single still image at a time.

- **Motion blur** is an important factor for achieving a natural, cinematic look. It is affected by the shutter speed, which should typically be set at double the frame rate (e.g., 1/50s for a 24fps video). Photography, on the other hand, often aims to freeze motion with faster shutter speeds.

- Videography often involves **recording synchronized audio** along with the visuals, necessitating the use of microphones and audio recording equipment. In photography, audio is not a factor, as the focus is solely on capturing still images.

- Videography requires attention to **continuity** across multiple shots and scenes to maintain a cohesive narrative. This may involve considerations like consistent lighting, wardrobe, and set design. In photography, while consistency may be important for a photo series, individual photos are often standalone pieces that don't require the same level of continuity.

- Videography requires the use of specific **video codecs and compression** techniques to reduce file size while maintaining quality. These are not a concern in photography, which typically uses formats like JPEG, TIFF, or RAW for still images.

- Videography involves a more complex post-production process, including **video editing**, **color grading**, **sound design**, and **visual effects**. While photography also involves post-production (image editing and retouching), the process is usually less time-consuming and intricate compared to videography.

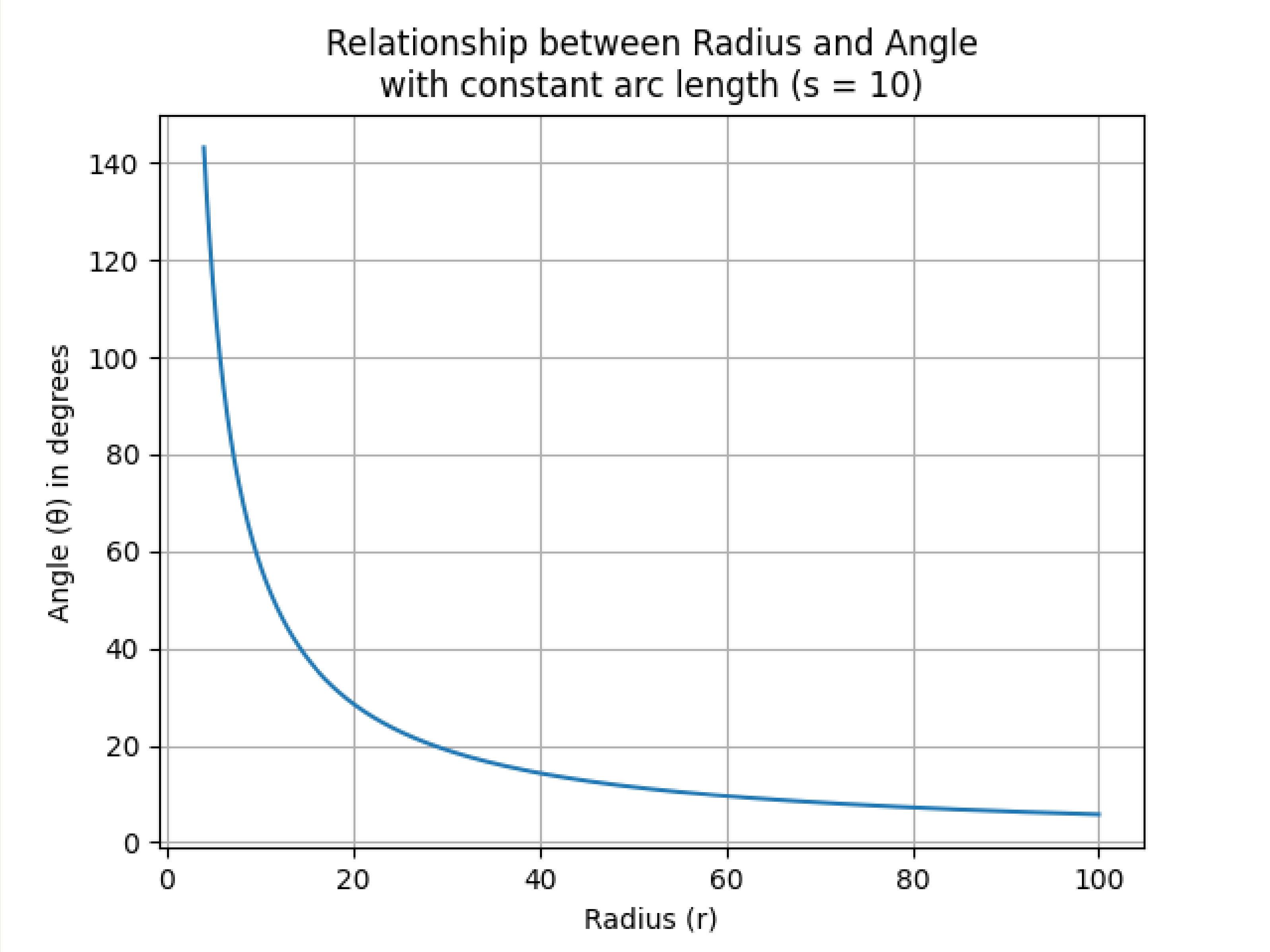
Perspective vs. Orthographic

Perspective projection is the type of image capture used by the eye, pinhole cameras, and lens cameras. Although the depth (radial) information is thrown away by definition, some clues to depth remain (the farther an object is from the camera/eye, the smaller it gets, but at a diminishing rate).

If we consider horizontal deviations from center of field of view as degrees theta (and vertical as phi):

θ (or ϕ) (degrees) = (s / r) * (180 / π); s = arc length, r = radius

If there were an orthographic camera, there would be a "filter of tiny straws" so that each sensor pixel is forced to look "straight ahead" (perpendicular to sensor). This is the same principle as architectural drawings, and preserves the actual distances between things on a given 2D plane. All depth information is lost here (no clues).



Visual Field as Theta-Phi(-Gamma)

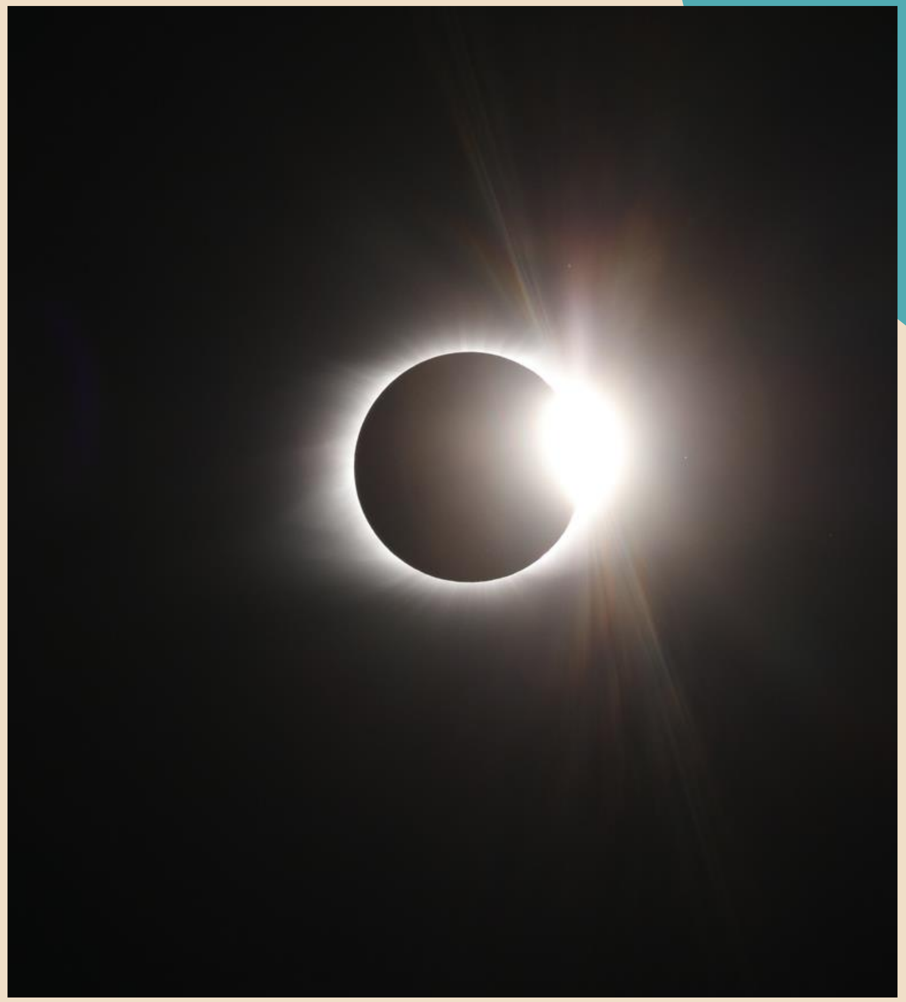
Although it's tempting to envision your visual field as a flat image with horizontal and vertical pixels, it's more helpful to think of it as comprising vector directions. The pupil serves as the convergence point, with the central foveal region being a vector aimed at your focus. Other pixels are accessed by deviating a specific number of theta or phi degrees from this central vector. Gamma represents the degrees from a vector extending from your right ear up to 90 degrees for the vector passing straight through the top of your head. Theta refers to horizontal deviations in vector direction, while phi denotes vertical deviations.

When using their fovea, humans can distinguish changes as small as 1/60th of a degree in both theta and phi, as well as similar amounts of depth. This equates to 50 million pixels across a typical field of view or 200 million pixels covering the entire surrounding sphere. Consequently, the number of perceivable voxels could potentially exceed **7 trillion**! Although a circle as **360** *degrees*, a sphere has **41,253** *square degrees*.

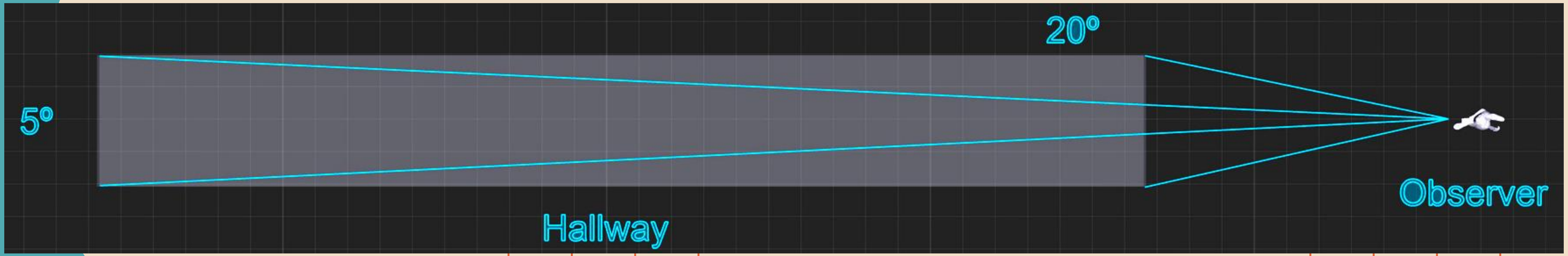
Visual Field as Theta-Phi(-Gamma)

The sun and moon each subtend 0.6 degrees (theta or phi), so roughly 1 degree. Your pinky fingernail, when held at arm's length, is roughly 2 degrees wide (theta) and tall (phi). Any object can be measured in terms of how much space it takes up on the theta-phi visual field. As you drive past a speed limit sign, it first grows to some maximal phi height, then shrinks again as you pass it, even though its actual height is constant.

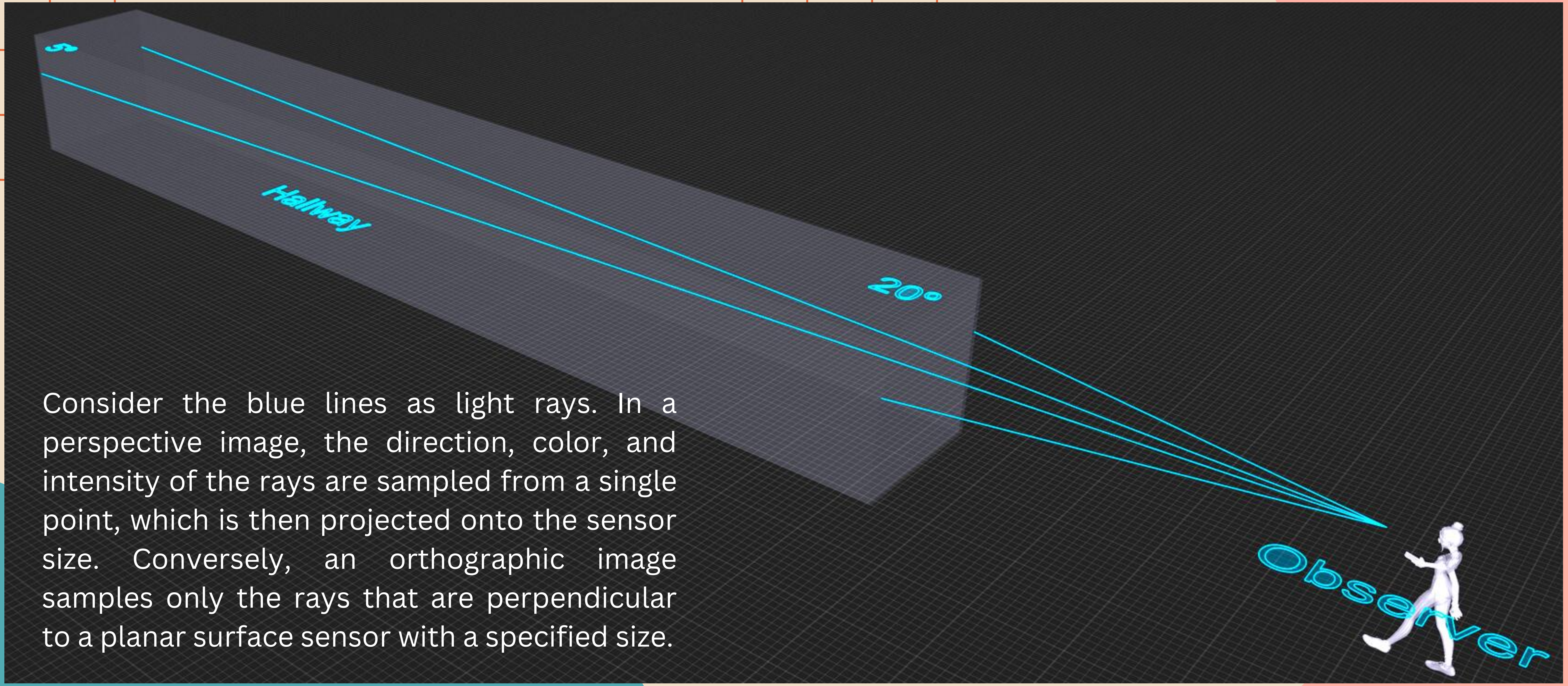
It is convenient that the sun and moon are roughly equal in theta size as viewed from earth, as it enables to seeing of a solar eclipse, as imaged here from 2017. Another eclipse will occur in 2024, going from Texas across to Maine, and having a longer duration of totality than 2017's.



Perspective vs. Orthographic



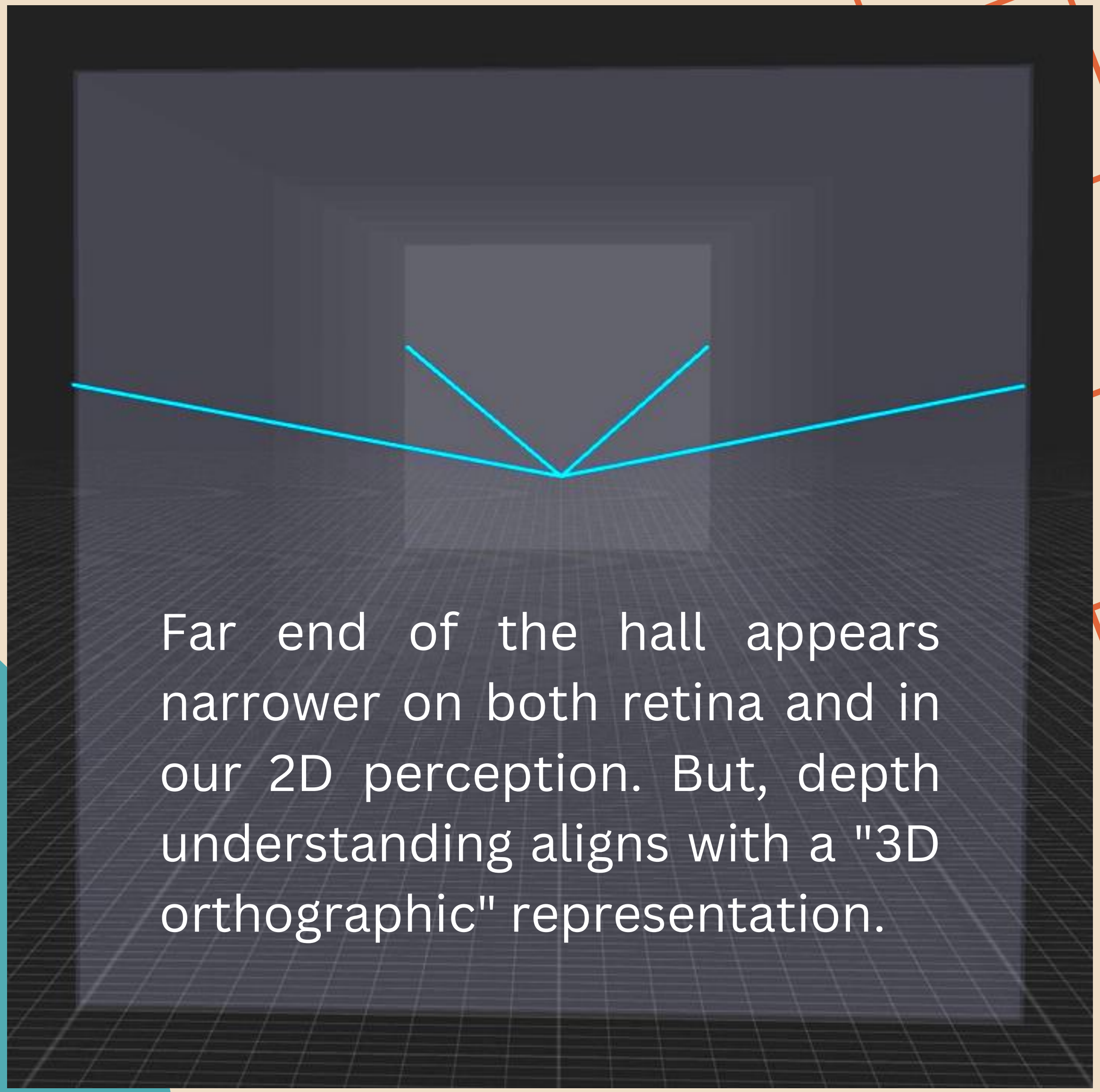
Orthographic (bird's eye view) depiction of why, in **girl's perspective view**, farther objects are shrunk as a matter of geometrical requirement.



Perspective view depiction of the same concept. Both ends of the hallway are 8 feet wide, but the front end of the hall takes up 20° of horizontal "theta" on girl's retina, whereas the back end of the hall takes up only 5° of theta.

Consider the blue lines as light rays. In a perspective image, the direction, color, and intensity of the rays are sampled from a single point, which is then projected onto the sensor size. Conversely, an orthographic image samples only the rays that are perpendicular to a planar surface sensor with a specified size.

Perspective vs. Orthographic



Girl's **perspective** view

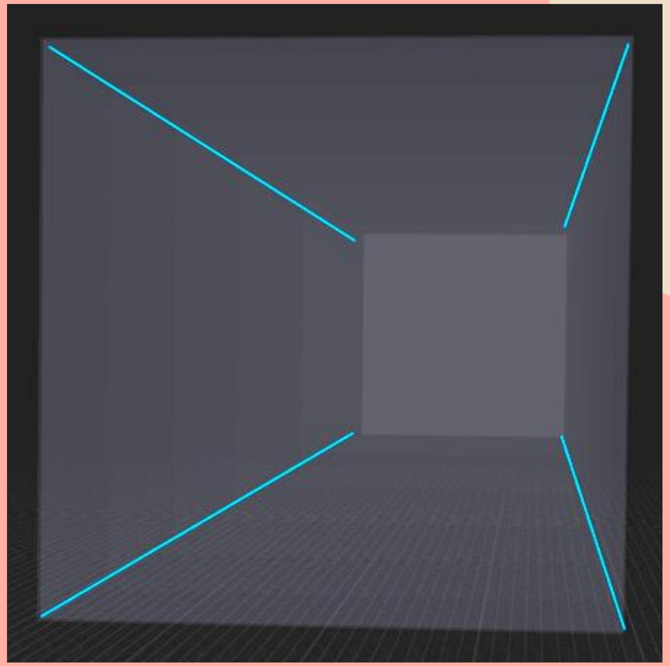
Orthographic view

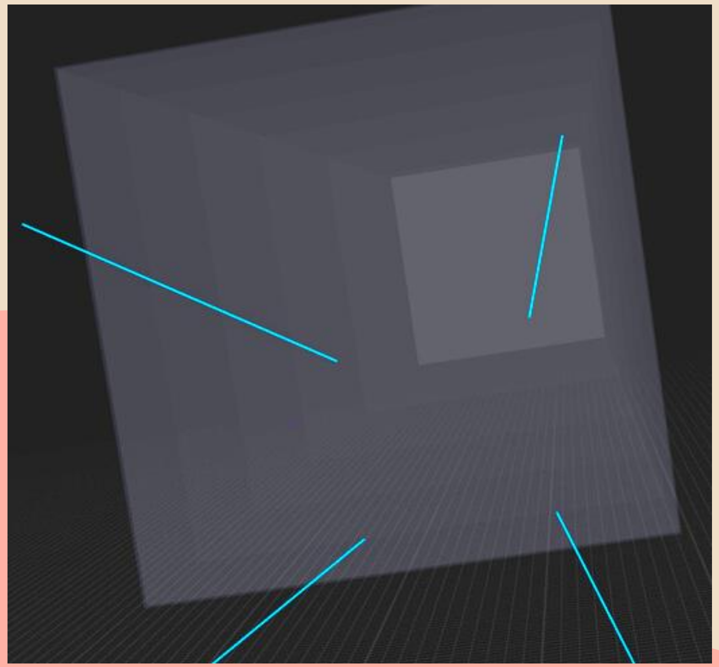
In the orthographic view, the front and back ends of the hall get superimposed over each other, since they are equally wide. Here, the sensor of the orthographic camera placed just behind the girl is 8 foot by 8 foot!

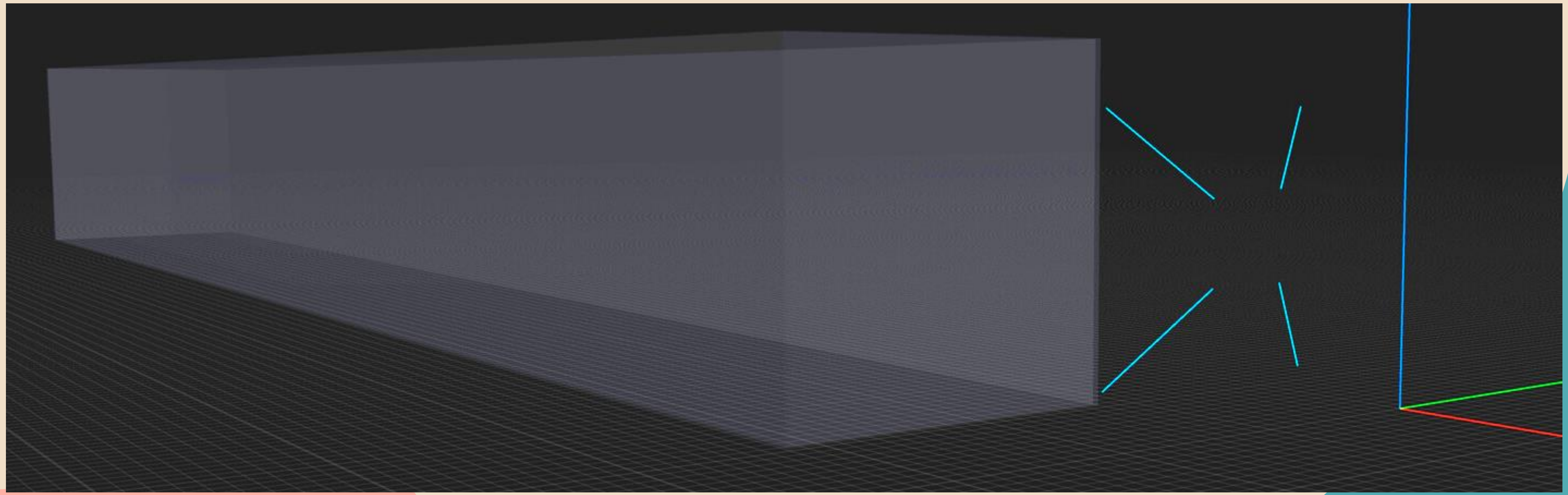
3D Orthographic, or Isometric

In the physical world, an object, like a 3-foot yardstick, maintains its dimensions regardless of its position. Similarly, within our brain's perception, objects preserve their proportions even when their 2D perspective appears smaller due to distance. Our mind interprets a 3D isometric model of the world based on the 2D perspective images we perceive.

In the example below, the computer program incorrectly interprets the blue lines, which we perceive as extending along the hallway, as being confined to the frontoparallel plane. Yet, our perception is something of a hybrid, as the lines going down the hall should be parallel.







Allocentric vs. Egocentric

Allocentric should be defined as the consistent 3D isometric representation of the environment, such as a hallway, with accurate dimensions. Egocentric, on the other hand, refers to the 2D perspective image derived from the 3D model, which inherently depends on a specific viewpoint in the 3D perceptual space and indicates the observer's location. This distinction holds true for mental imagery, where a 3D mental space (allocentric) and a 2D perspective image (egocentric) serve to position the mind's eye within that space.

The uniform 3D isometric depiction of reality or imagination forms a "skin" or schema.

Allocentric ≈ **Isometric** ≈ **Orthographic** (3D "Boundaries/Skins" or 2D "Edges/Perimeters")

Egocentric ≈ **Perspective** (usually 2D)