Welcome to Computer Vision

Computer Vision

Dr. Muhammad Tahir

DEPARTMENT OF COMPUTER SCIENCE,

FAST-NUCES, Peshawar

Course Details

LECTURES: Monday & Wednesday

TIMINGS:

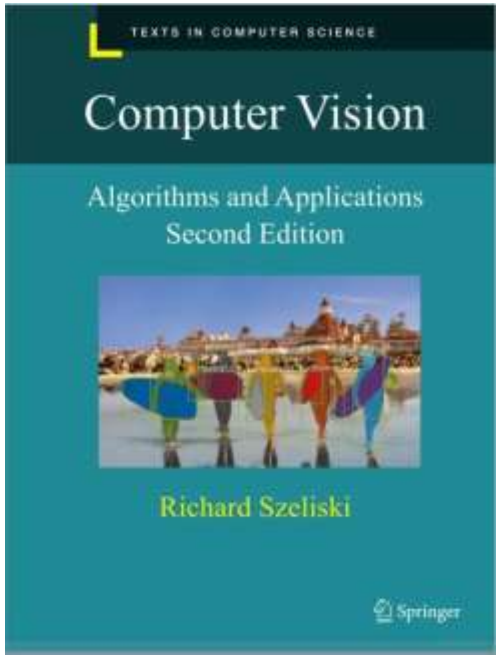
9:30 am – 11:00 am

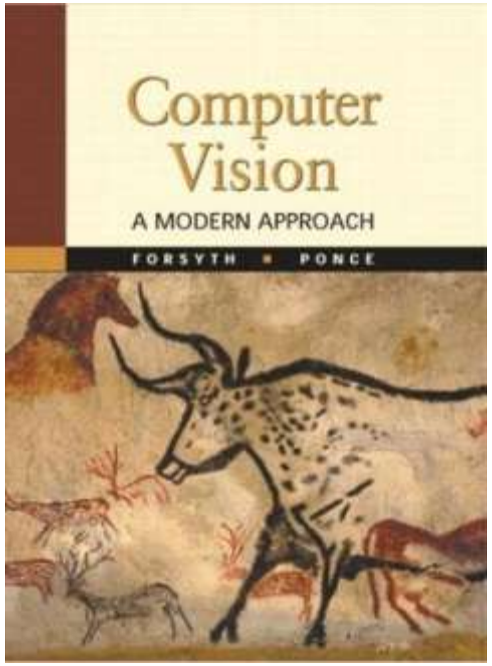
MY OFFICE:

OFFICE HOURS:

EMAIL: m.tahir@nu.edu.pk

References



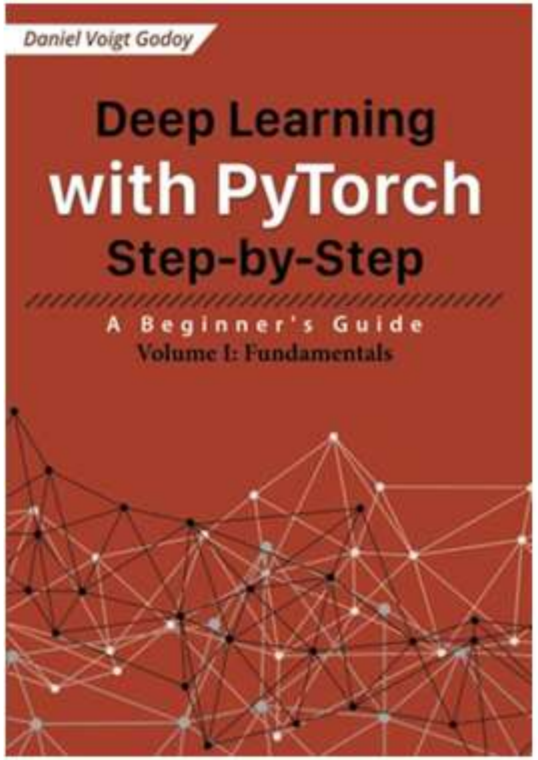


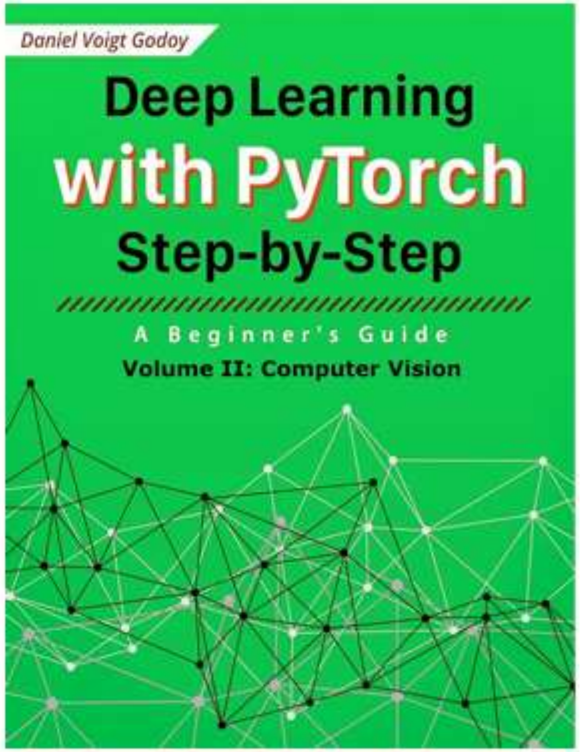
The material in these slides are based on:

1. Rick Szeliski's book: [Computer Vision: Algorithms and Applications]()
2. Forsythe and Ponce: [Computer Vision: A Modern Approach]()

Recommended Books

Deep Learning with PyTorch Step-by-Step by Daniel Voigt Godoy





Course Learning Outcomes

No	CLO (Tentative)	Domain	Taxonomy Level	PLO
1	Understanding basics of Computer Vision: algorithms, tools, and techniques	Cognitive	2	
2	Develop solutions for image/video understanding and recognition	Cognitive	3	
3	Design solutions to solve practical Computer Vision problems	Cognitive	3	

Outline

Image Formation

Image Formation

- **Image Formation** is the Projection of 3D scene onto 2D plane.
- We need to understand the **geometric and photometric relation between the scene and its image**.
- By **geometric** we mean, given a point in the scene we want to understand where it ends up in the image.
- By **photometric** we mean, given the brightness and appearance of a point in this scene we want to understand what the brightness and appearance would be in the image.

Pinhole and Perspective Projection

- So, we are going to start with the concept of a **pinhole camera**.
- This is the simplest type of camera you can imagine.
- The pinhole camera performs what's called **perspective projection**.
- This is one of the most important concepts in computer vision.
- We will derive perspective projection and talk about some of the manifestations of perspective projection.
- Then we will argue that while the pinhole camera is great in terms of the clarity of images it can produce, it simply does not gather enough light.
- To resolve this issue, we use **lenses**.

Image Formation Using Lenses

- We will talk in great depth about image formation using lenses
- Lenses of various attributes, focal length, depth of field, defocus, F number
- We will discuss all of these different characteristics of lenses.

Lense Related Issues

- Then we will talk about various issues related to lenses
- Even if a Lense is perfectly manufactured, it is going to end up having some geometric aberrations and produce some distortions.
- We will talk about what these effects are in images and how we might be able to correct them.

Wide Angle Cameras

- Next, we will deviate from perspective projection and look at the problem of capturing unusually large fields of view
- For example, hemispherical field of view
- This cannot be done using perspective projection
- So, we design lenses, which will allow us to capture very large fields of view, as well as combinations of mirrors and lenses.
- And finally, we will talk about biological eyes
- We will talk about some fascinating designs that nature has come up with.
- Then we will focus on the human eye and some of its remarkable characteristics.

Pinhole and Perspective Projection

Image Formation

- So, here you see a 3D scene on the right, which is a house.
- And you see a screen, or let's call it an image plane on the left.
- The question that I would like to pose to you:

Is an image of the house being formed on the screen?

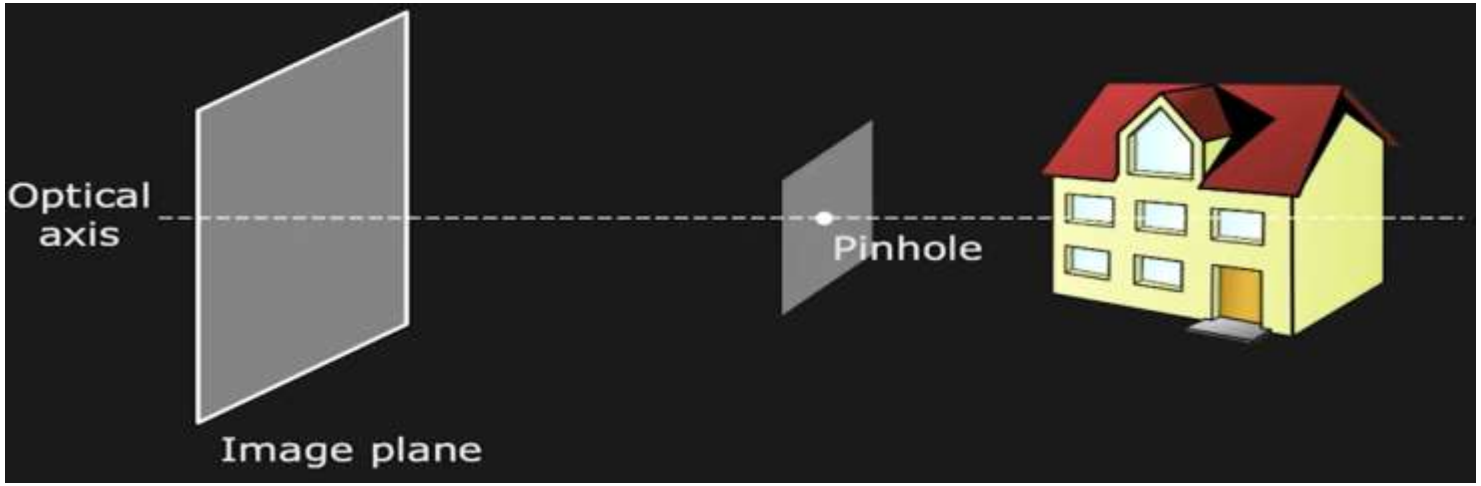
Image Formation

If you consider any point on the screen, it does receive light from a lot of points on the house, but you don’t see a clear image.

So, one could argue that there is an image being formed, but it is a muddled image and not a clear one.

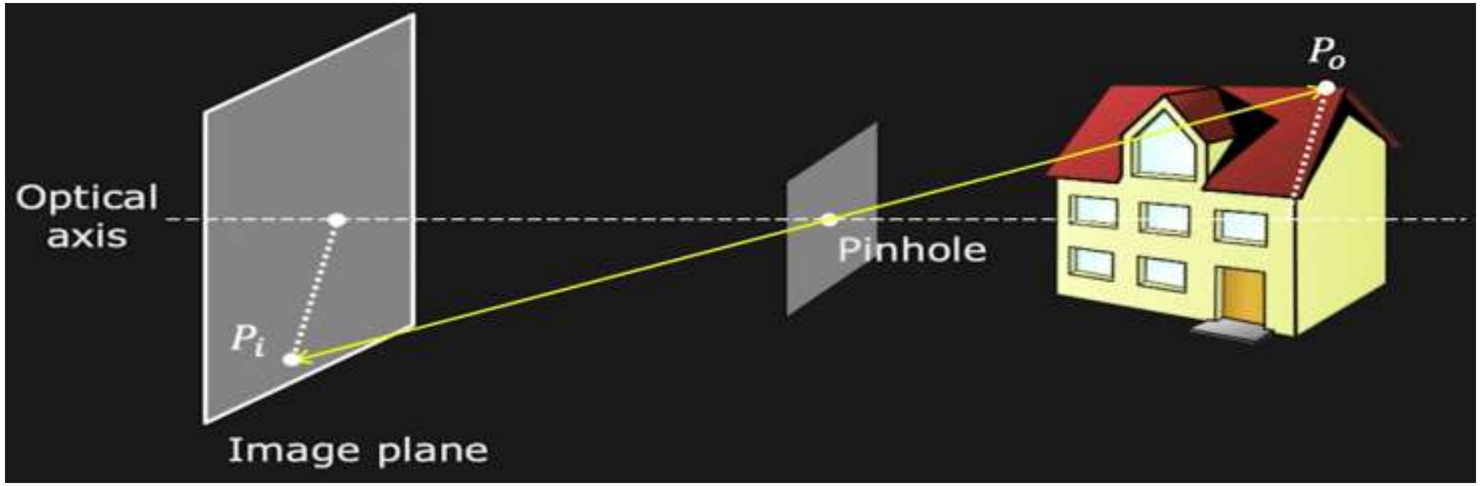
So, then the question is, how does one create a clear, crisp image of the house on the screen.

Perspective Imaging with Pinhole



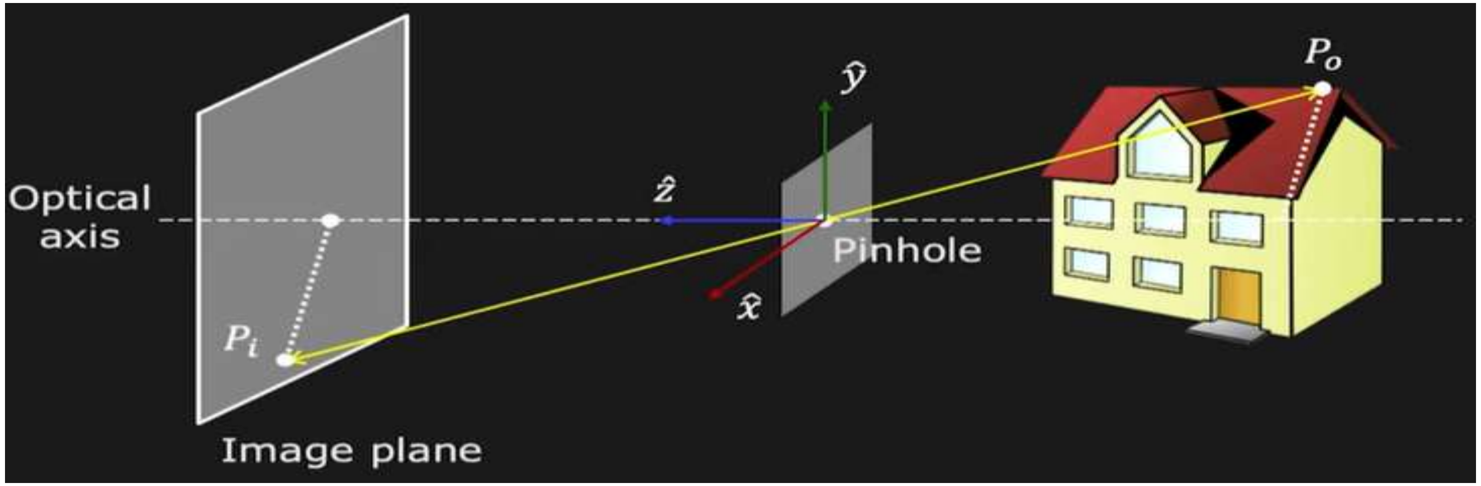
- The simplest way to do this is by using a pinhole.
- A pinhole is an opaque sheet with a tiny hole in it and it is placed between the scene and the image plane.

Perspective Imaging with Pinhole



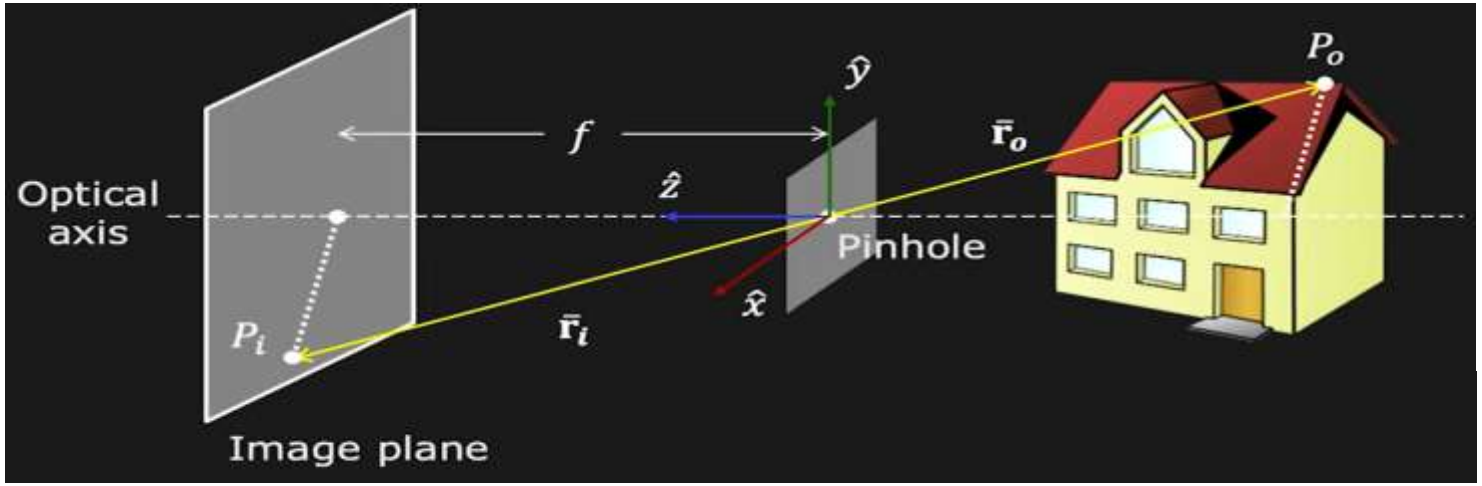
- If you take a look at a single point on the house, you see that there is a single ray that travels from that point, in this case, P_0, to the image plane and projects onto the point P_i.
- So, every point of the scene now projects onto a single point in the image.

Perspective Imaging with Pinhole



- We want to understand the relationship between P_0 and P_i.
- First, we are going to erect a coordinate frame xyz, 3D coordinate frame placed at the pinhole, with the $z - axis$ pointing along the optical axis.
- The optical axis is the axis that is perpendicular to the image plane i.e. the dotted line in this image.

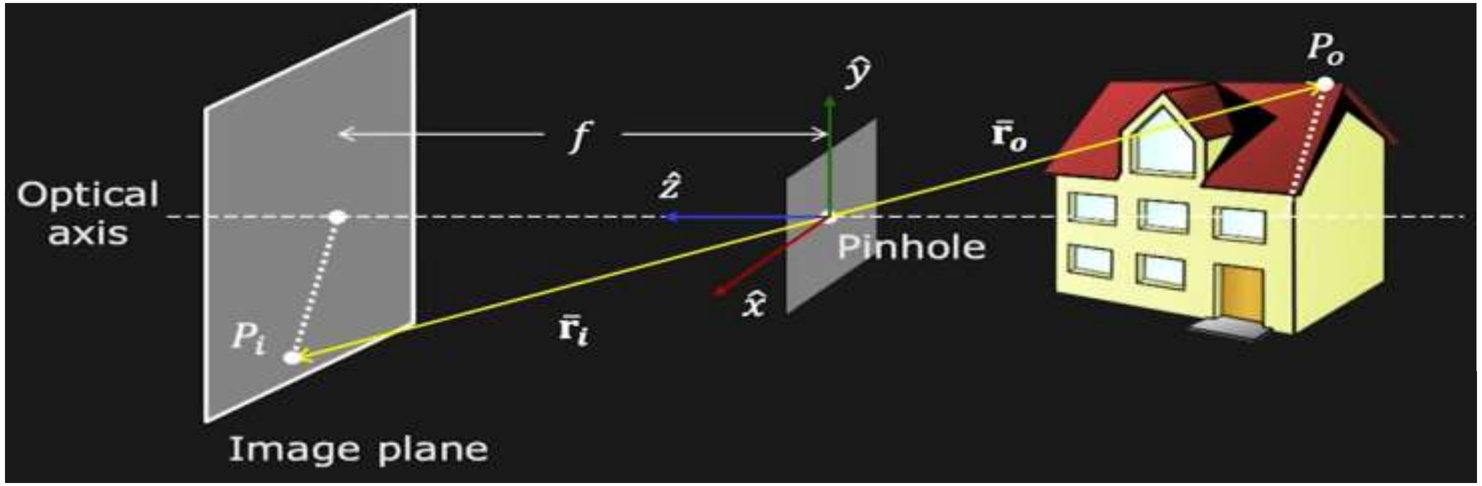
Perspective Imaging with Pinhole



- The distance between the pinhole and the image plane is called the effective focal length f.

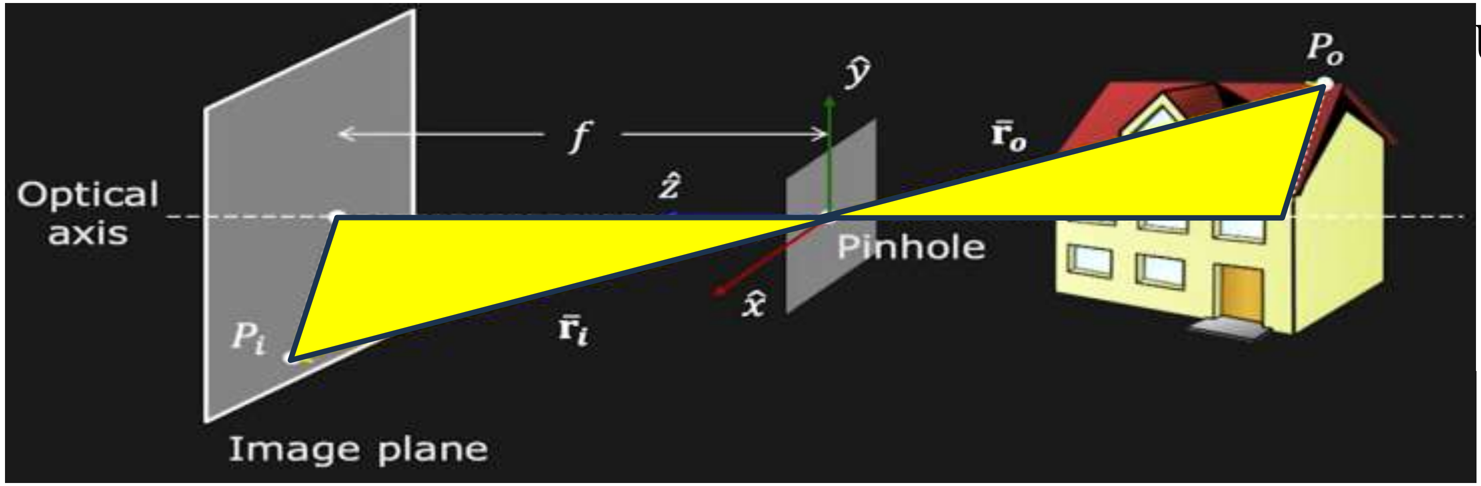
$f: Effective\ Focal\ Length$
$\bar{r}_0 = (x_0, y_0, z_0)$
$\bar{r}_i = (x_i, y_i, f)$

Perspective Imaging with Pinhole



- Now, we can write the point using P_0 using the vector r_0, which has coordinates (x_0, y_0, z_0)
- Its image can be denoted as r_i, which has coordinates x_i on the image plane, y_i, and the $z - coordinate$ is going to be f.
- Irrespective of where the point lies in 3D scene, irrespective of its z component, the z component on the image plane is always going to be f.

Perspective Imaging with Pinhole

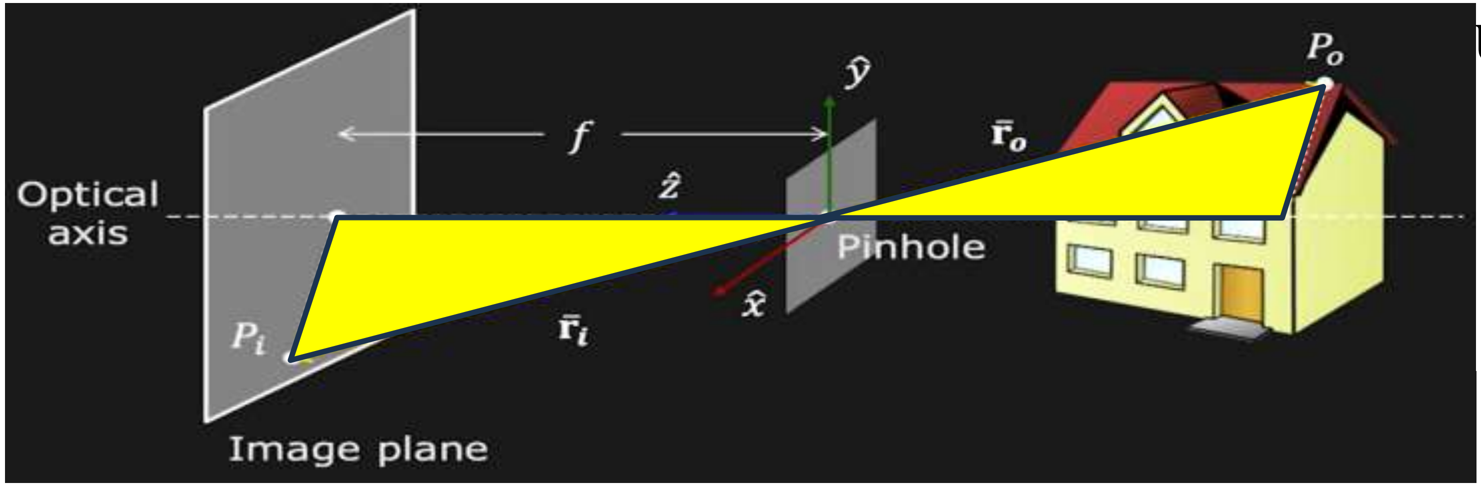


Using Similar Triangles

$$\frac{\bar{r}_i}{f} = \frac{\bar{r}_0}{z_0}$$

- If you consider the two yellow triangles, these are similar triangles
- From these similar triangles, you can write the vector r_i divided by f is equal to the vector r_0 divided by z_0
- z_0 is the z component of the 3D point. In other words, it is the depth of the point in 3D.

Perspective Imaging with Pinhole



Using Similar Triangles

$$\frac{\bar{r}_i}{f} = \frac{\bar{r}_0}{z_0}$$

- Since r_i and r_0 are vectors, we can break it down into its components

$$\frac{\bar{r}_i}{f} = \frac{\bar{r}_0}{z_0} \quad \rightarrow \quad \frac{x_i}{f} = \frac{x_0}{z_0}, \frac{y_i}{f} = \frac{y_0}{z_0}$$

- These are the equations of perspective projection.

$$\frac{\bar{r}_i}{f} = \frac{\bar{r}_0}{z_0} \quad \rightarrow \quad \frac{x_i}{f} = \frac{x_0}{z_0}, \frac{y_i}{f} = \frac{y_0}{z_0}$$

- They are very simple equations but creates some non-intuitive effects in the image.
- This idea of pinhole projection and pinhole camera actually dates back to 500 BC.

Camera Obscura – Dark Chamber

- There were Chinese philosophers who were writing about this concept
- In 1000 AD, the Arab Physicist, Alhazen, he wrote a book Kitab Al-Manazir where he described the concept of pinhole camera in great details.
- This concept came to the West in 16^{th} century
- This became very popular among artists as a tool for rendering accurate depictions of 3D scenes.

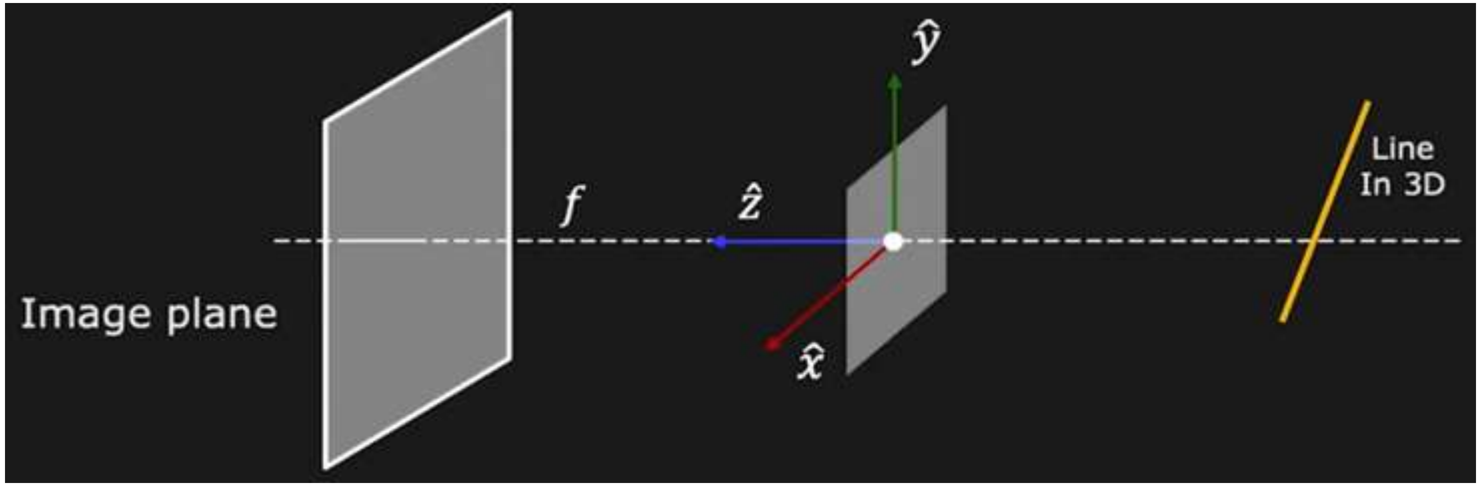
Camera Obscura – Dark Chamber

- This is a sketch by Gemma Frisius, the Dutch mathematician, where you can see that there is a wall with a little pinhole and behind this wall is a second wall.
- So, the 3D scene is projected by this pinhole onto a 2D image
- Now you can imagine, that an artist can walk up to this wall and sketch out a very accurate geometrical representation.
- This concept is called Camera Obscura which in Latin means a Dark Chamber.

Properties of Perspective Projection

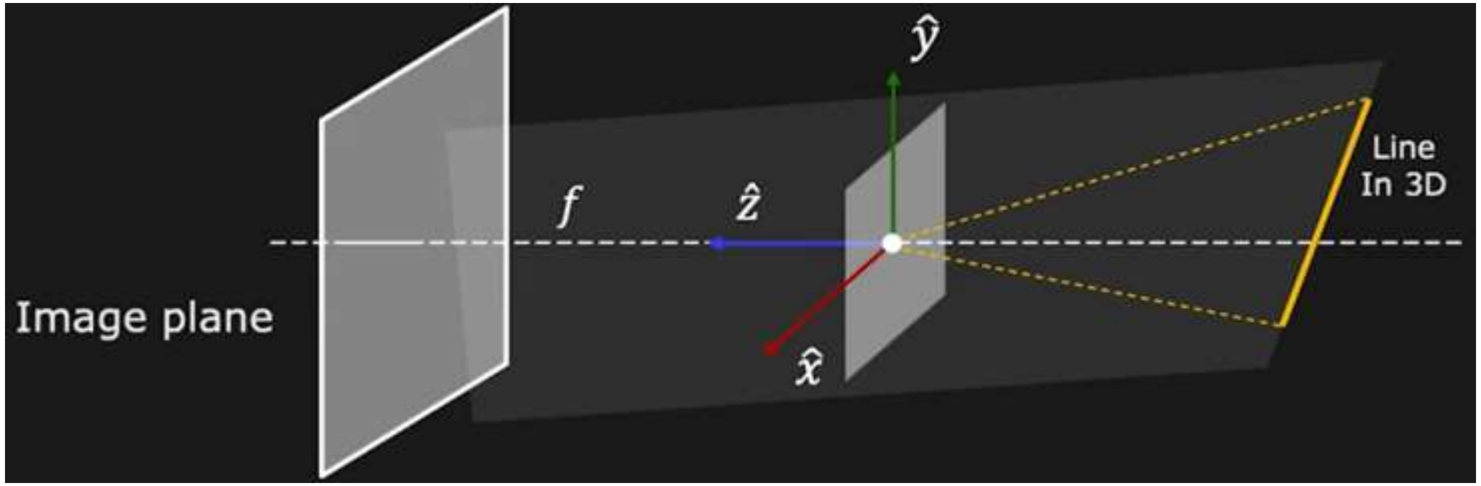
Perspective Projection of a Line

Perspective Projection of a Line



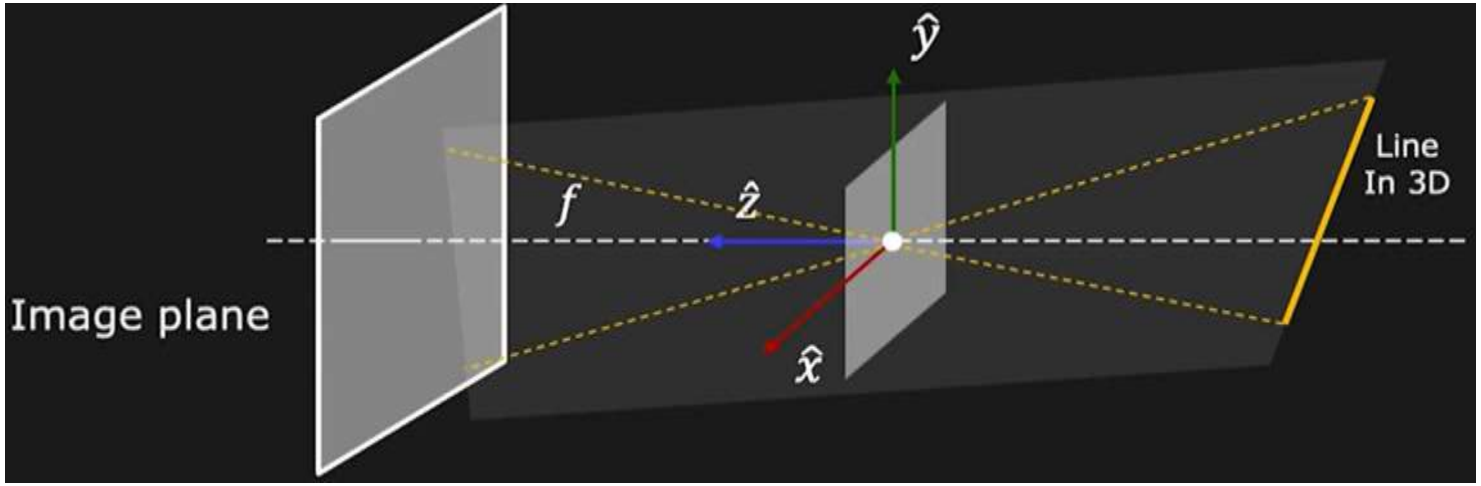
What is a perspective projection of line in 3D?

Perspective Projection of a Line



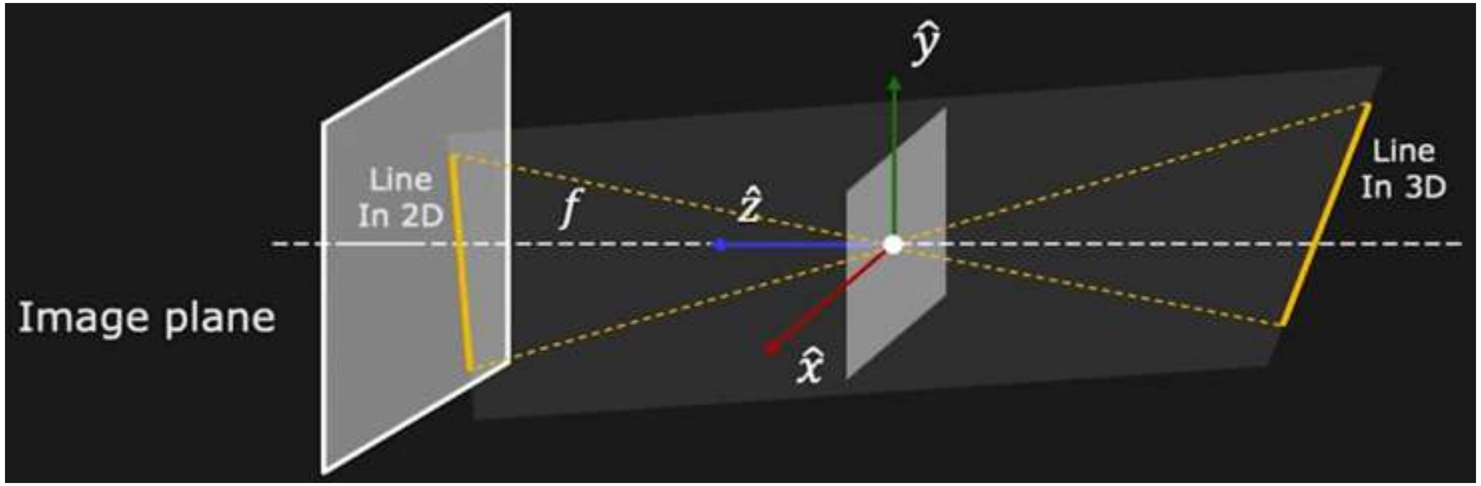
- The line and a pinhole (i.e. a point here) define a plane in 3D.
- All the rays of light that pass through the pinhole lie on this plane.

Perspective Projection of a Line



- Therefore, all the light that pass through the pinhole towards the image plane almost also should lie on this plane.

Perspective Projection of a Line



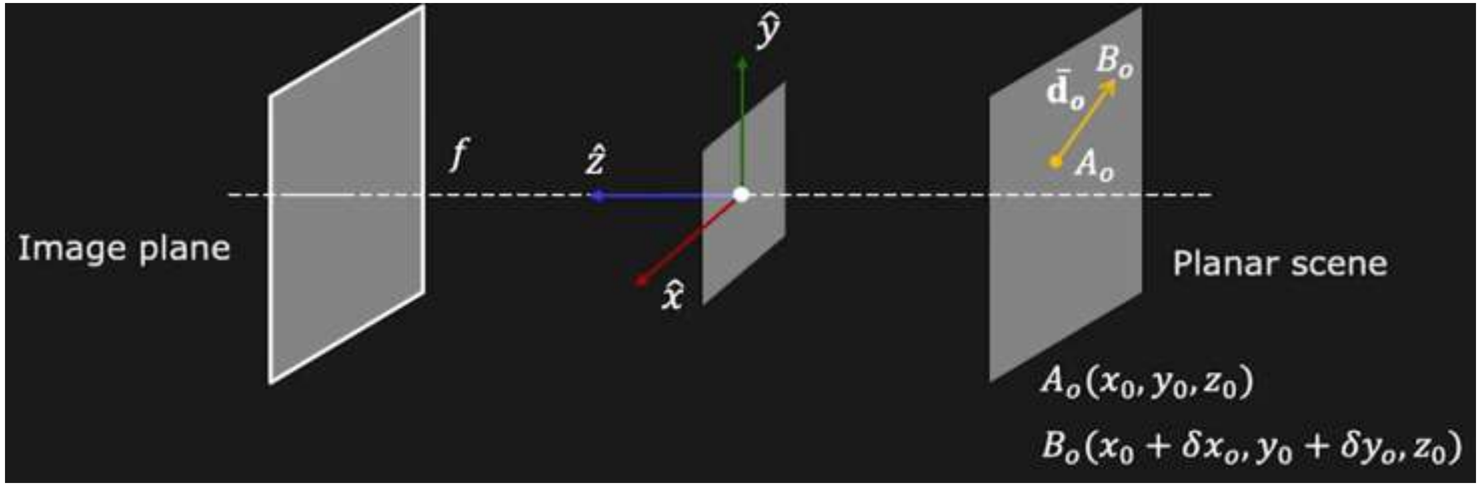
- So, the image of this 3D line on your 2D image must lie at the intersection of this plane and the image plane.

Perspective Projection of a Line

- The image of a line in 3D has to be a line in 2D
- That is what we observe in images, straight line in scene remains straight in image.

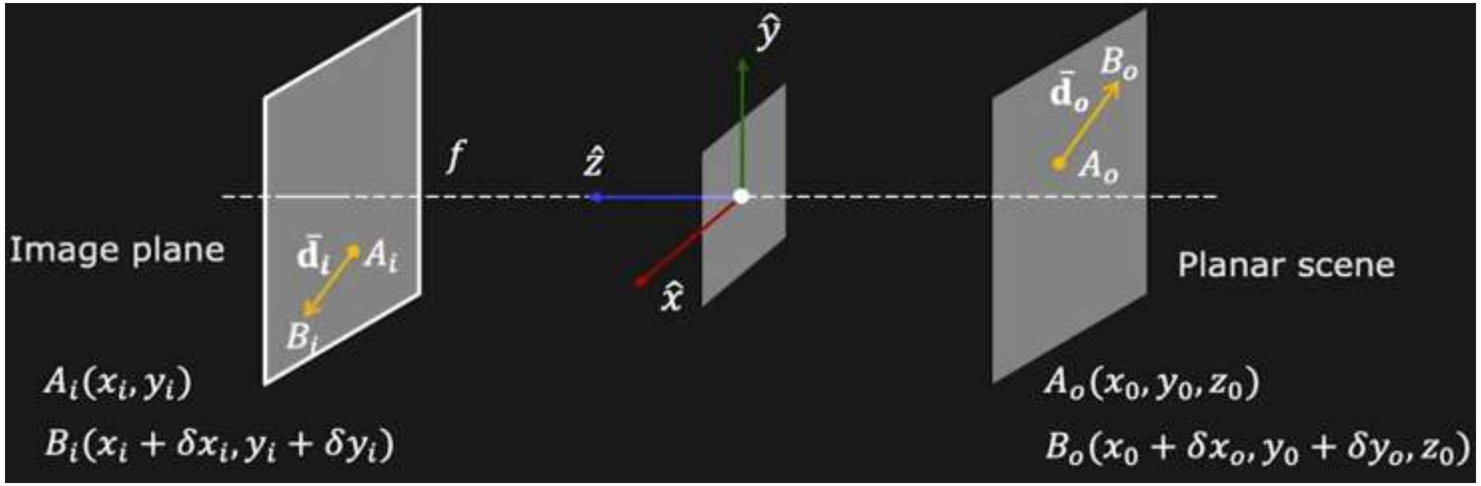
Image Magnification

Image Magnification



- An object with certain size at a certain distance, what is going to be its size in the image?
- Let's consider a little segment here A_0, B_0 of length d_0
- This segment lies on the scene that is parallel to the image plane

Image Magnification



- It produces an image which is another segment and that segment is A_i, B_i
- We need to know the length of the segment A_i, B_i due to a segment of length d_0 in the scene.

Image Magnification

- The ratio of the length of the segment in the image to the length of the segment in the scene is called the magnification.

$$\text{Magnification: } |m| = \frac{\|\bar{d}_i\|}{\|\bar{d}_0\|} = \frac{\sqrt{{\delta x_i}^2 + {\delta y_i}^2}}{\sqrt{{\delta x_0}^2 + {\delta y_0}^2}}$$

From Perspective Projection,

$\frac{x_i}{f} = \frac{x_0}{z_0}$ and $\frac{y_i}{f} = \frac{y_0}{z_0}$ $\longrightarrow$ **(A)**

$\frac{x_i + \delta x_i}{f} = \frac{x_0 + \delta x_0}{z_0}$ and $\frac{y_i + \delta y_i}{f} = \frac{y_0 + \delta y_0}{z_0}$ $\longrightarrow$ **(B)**

Image Magnification

- From (A) and (B), we get:

$$\frac{\delta x_i}{f} = \frac{\delta x_0}{z_0} \text{ and } \frac{\delta y_i}{f} = \frac{\delta y_0}{z_0}$$

- This shows the relationships between the displacements in the image to the displacements in the scene.
- Put this back into the magnification equation

$$|m| = \frac{\|\bar{d}_i\|}{\|\bar{d}_0\|} = \frac{\sqrt{\delta {x_i}^2 + \delta {y_i}^2}}{\sqrt{\delta {x_0}^2 + \delta {y_0}^2}} = \left|\frac{f}{z_0}\right|$$

Where f is the focal length and z_0 is the depth of the object in the scene.

Image Magnification

- $m = \frac{f}{z_0}$, m is negative when image is inverted. So, in case of pinhole camera, it will be negative.
- Note z_0 in the denominator
- The size of the magnification of an object in an image is inversely proportional to its distance from the camera.

Image Magnification

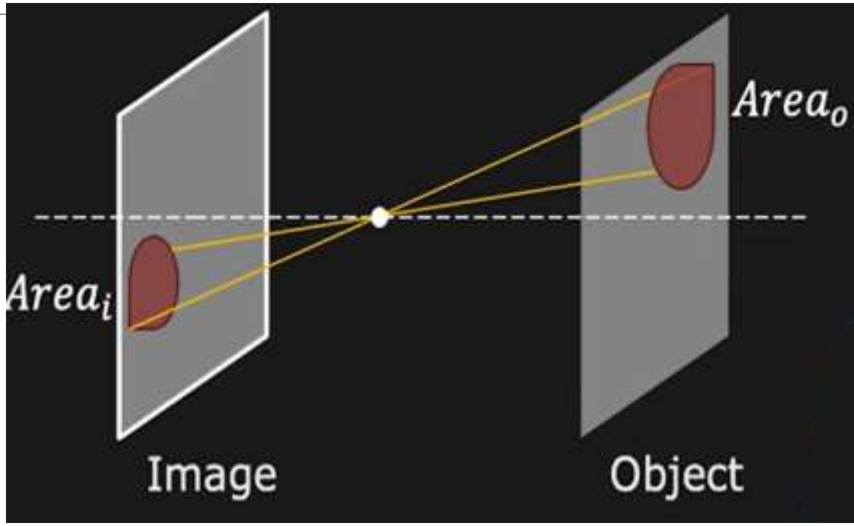
- We know that the two lines on a railway track are parallel in the 3D scene.
- But in the image, they appear to be intersecting at infinity.
- As you go further and further away in terms of depth, the two lines get closer and closer
- This is because Magnification is inversely proportional to the depth i.e. z_0 in $m = \frac{f}{z_0}$

Image Magnification - Remarks

- Magnification of an object can be assumed to be constant if the object is small compared to its distance from the camera.
- If you consider a small object, its range of depth values is small compared to its distance from the camera then it can be assumed that the entire object is subjected to the same magnification.

- m can be assumed to be constant if the range of scene depth Δz is much smaller than the average scene depth $\tilde{z}$

Image Magnification - Remarks



The ratio of the area of an object in the image to its area in the scene is $m\ squared$.

$$\frac{Area_i}{Area_0} = m^2$$

Example

The **object height** $h_o = 4\ cm$

The **image height** $h_i = 2\ cm$

Then the **magnification** is:

$$m = \frac{h_i}{h_o} = \frac{2}{4} = 0.5$$

Compute the area ratio

$$\frac{Area_i}{Area_o} = m^2 = (0.5)^2 = 0.25$$

So,

$$Area_i = 0.25 \times Area_o$$

Interpret

If the object had an **area of 40 cm²**, then

$$Area_i = 0.25 \times 40 = 10\ cm^2$$

Vanishing Point

Vanishing Point

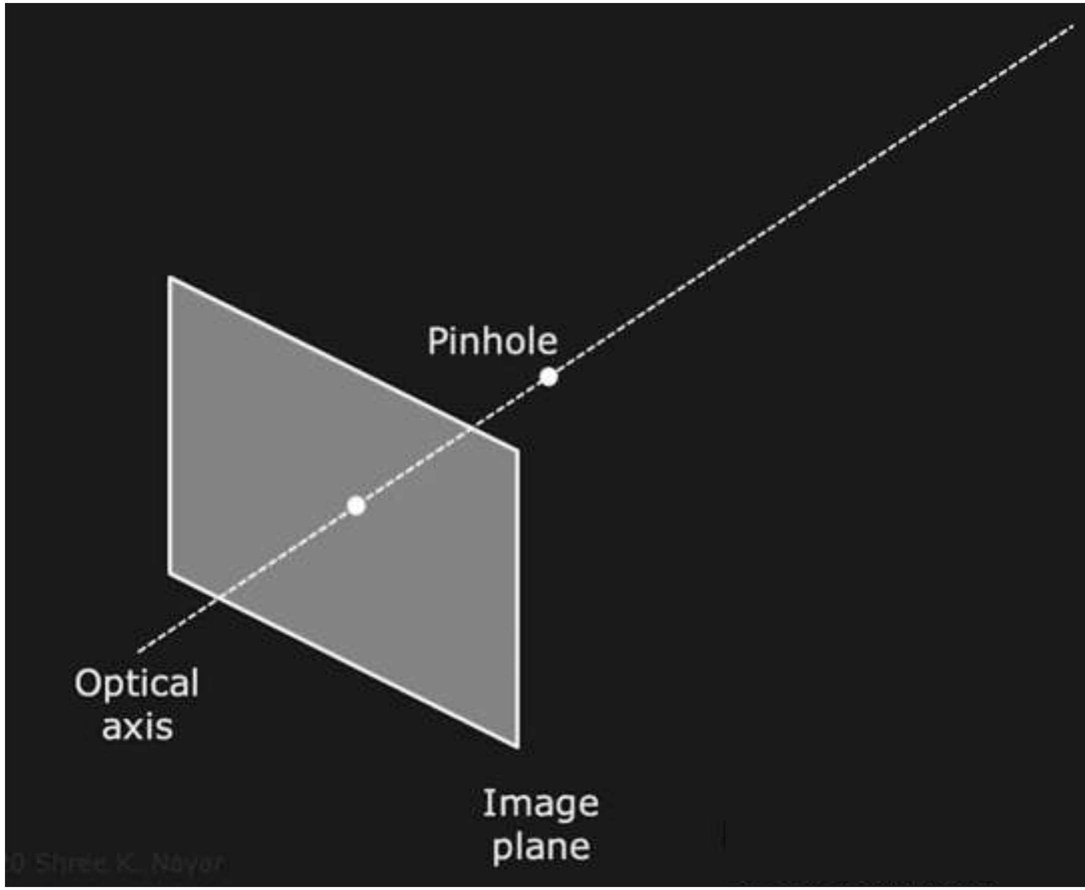
- There are lines in this image
- Consider the road lanes, the lights, the lines that separate the yellow wall and grey wall.
- We can assume that all these lines are parallel in 3D.
- All these lines seem to be emerging from a single point in the image and that point is called the vanishing point.

Vanishing Point

- Parallel straight lines converge at a single image point
- Location of this Vanishing Point in the image depends on the orientation of parallel straight lines in 3D.

Finding Vanishing Point

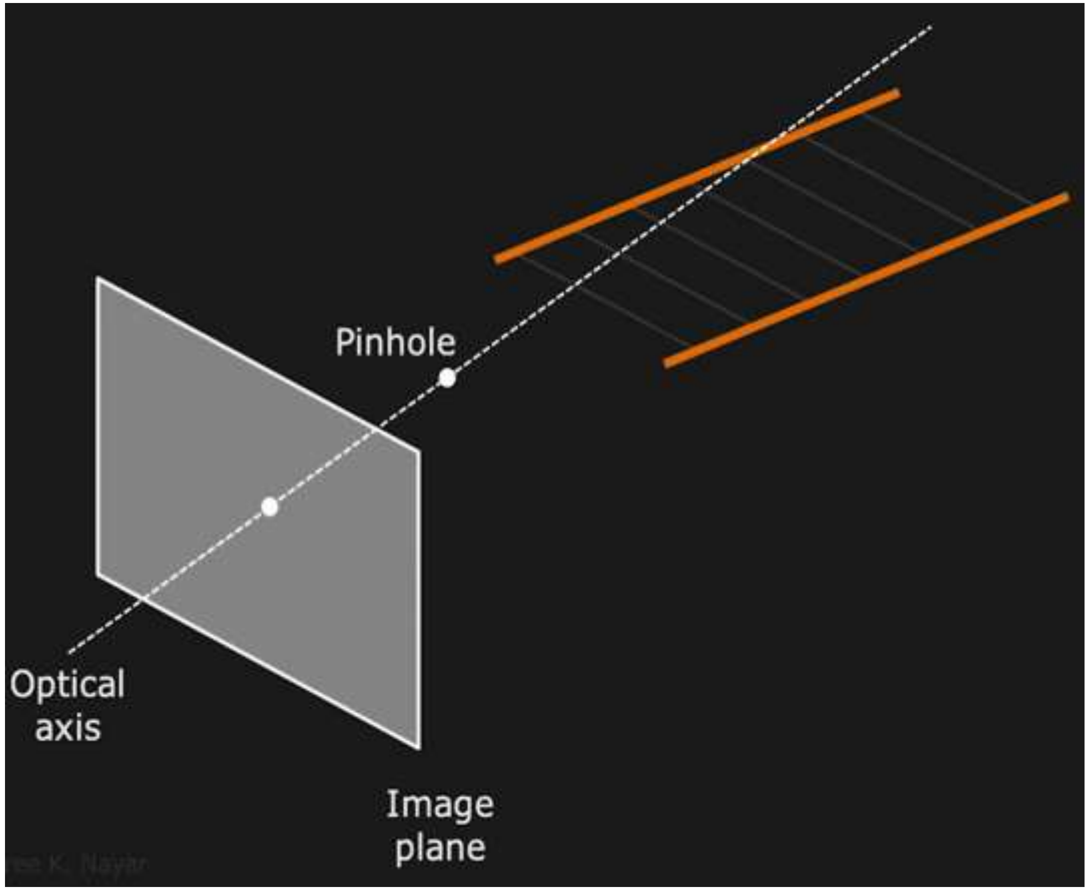
- So given a set of parallel lines in 3D, how to know where the vanishing point is going to end up for that set of lines.



Finding Vanishing Point

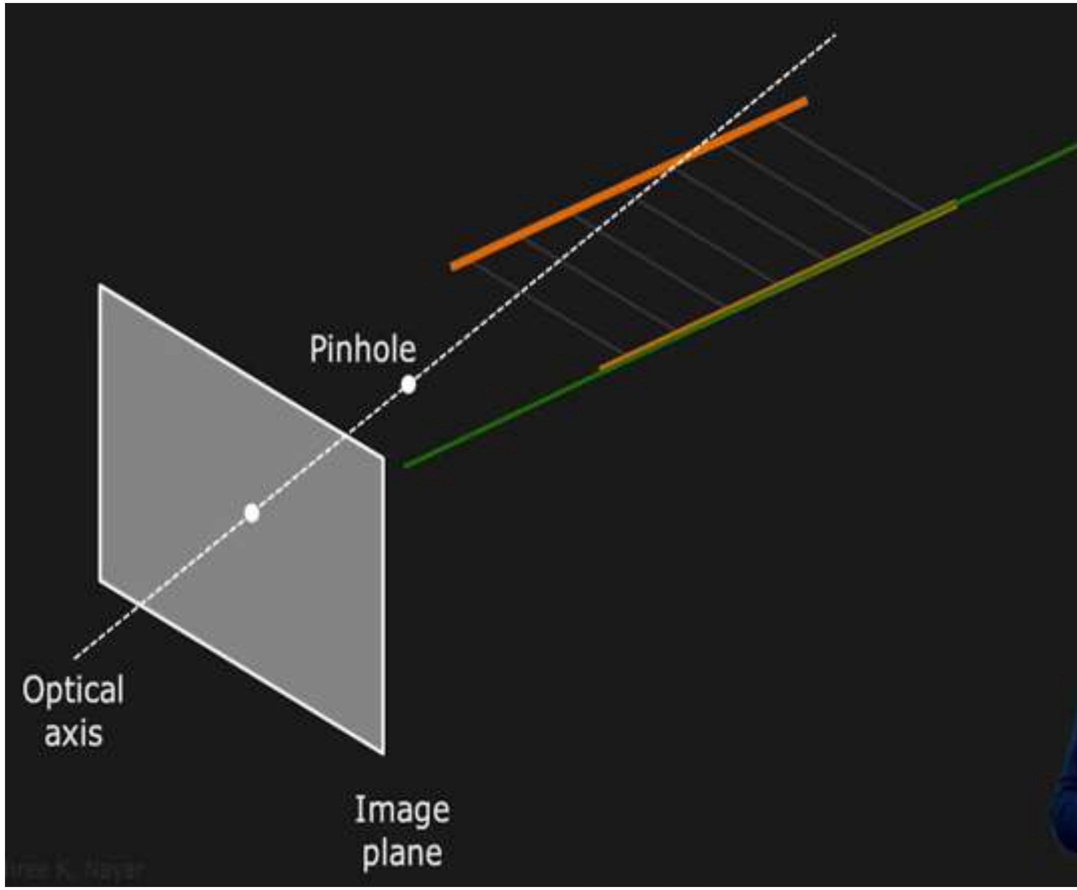
- Assume these two parallel lines and we want to find the vanishing point corresponding to these two lines.

- Note: All parallel lines in 3D share the same vanishing point.



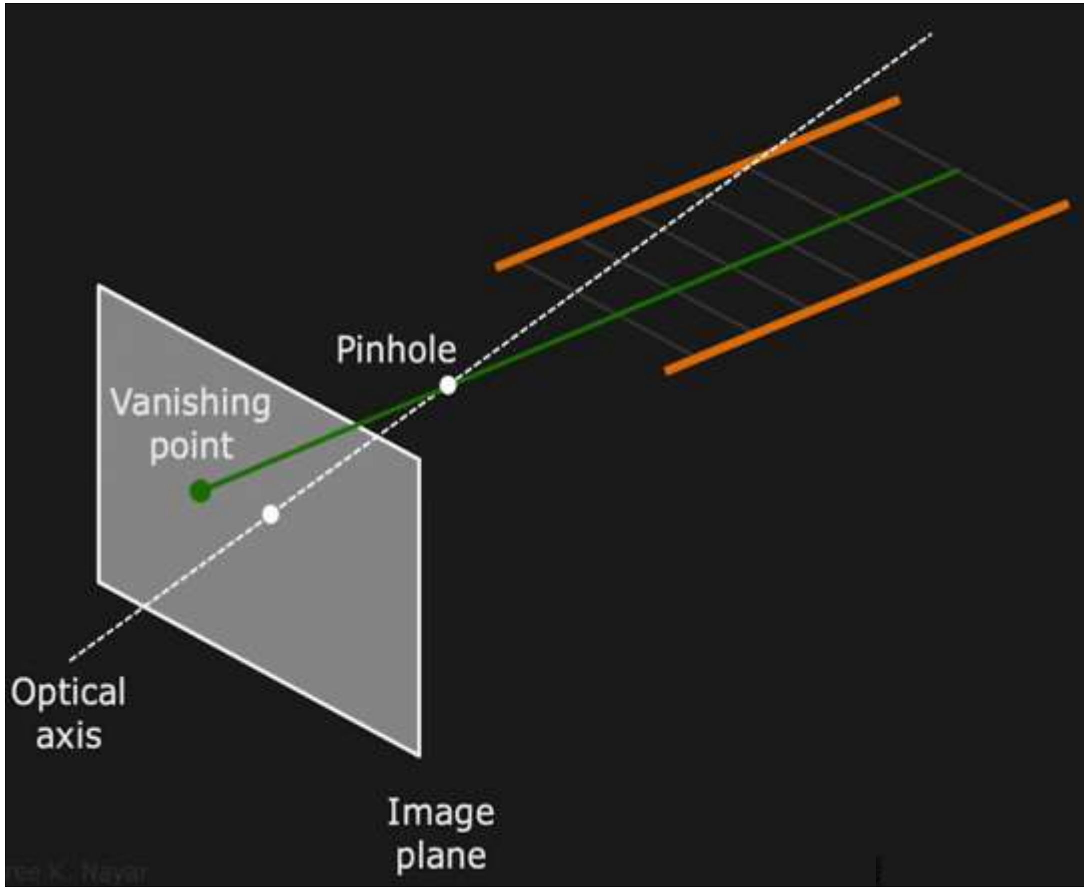
Finding Vanishing Point

So, construct a line (i.e. green line) that is parallel to these two lines that passes through the pinhole.



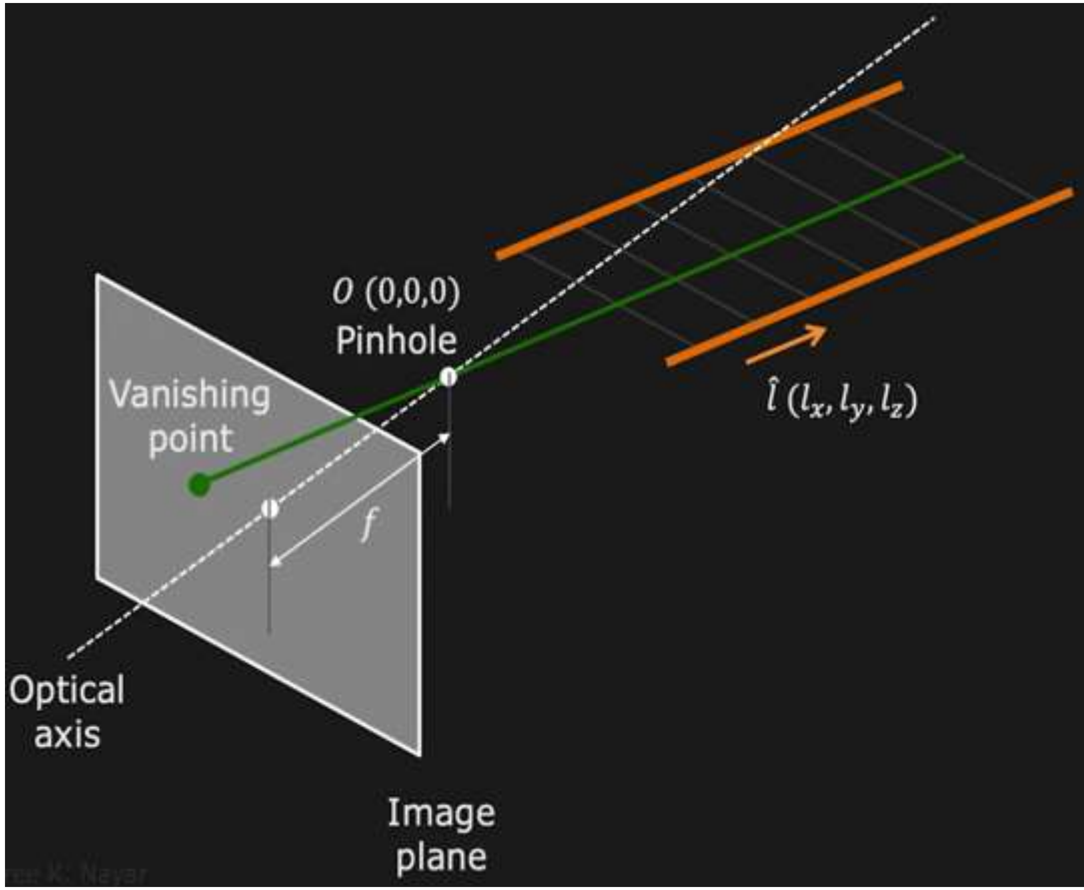
Finding Vanishing Point

- Wherever that line pierces the image is the vanishing point corresponding to this set of parallel lines in 3D.



Finding Vanishing Point

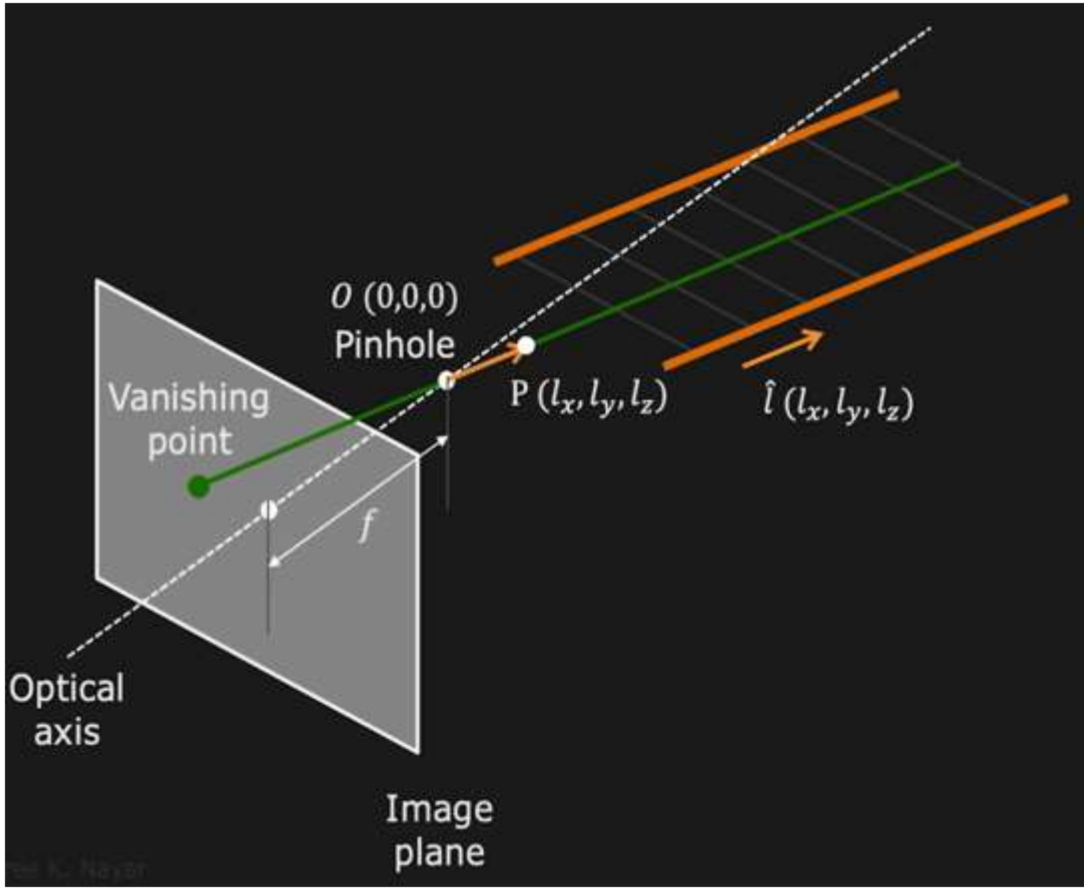
- Given that we know the perspective projection, we first define the direction of the set of parallel lines in 3D
- This is given by the vector l_x, l_y, l_z
- Then we create a point, which is in that direction.



Finding Vanishing Point

- Point P in that direction from the pinhole of the camera
- Perspectively project that point into the image using perspective projection equations.
- Vanishing point of the line is the projection of point P.

$$(x_{vp}, y_{vp}) = \left(f\frac{l_x}{l_z}, f\frac{l_y}{l_z}\right)$$



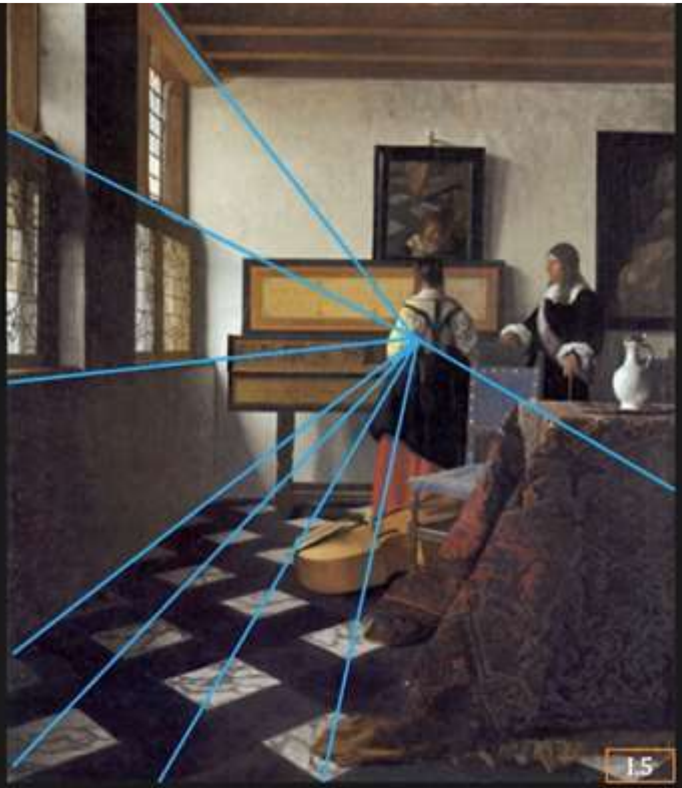
Vanishing Point

- The concept of vanishing point has been extensively used by artists.
- Johannes Vermeer painted The Music Lesson
- There are a lot of parallel lines or set of parallel lines

The Music Lesson, Johannes Vermeer, c. 1662-1664

Vanishing Point

- Let’s consider this set of parallel lines, which converge at the point shown (vanishing point)
- The artist of this painting has placed the most important object or activity in his opinion at this point.
- He wants to draw your attention to that activity.

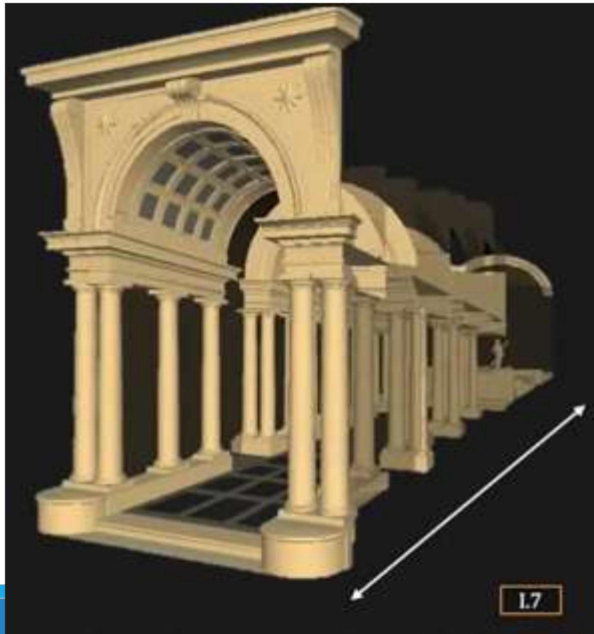


The Music Lesson, Johannes Vermeer, c. 1662-1664

False Perspective

- This is called Galleria Spada by Francesco Borromini, 1652, beautiful little gallery in Rome.
- What we see here is an archway or a little bit of a hallway and at the end of this hallway, there is a sculpture.
- When you stand in front of the hallway, you get the impression that the sculpture is roughly 150 feet away from you.
- But in reality, this sculpture is 30 feet away from you.

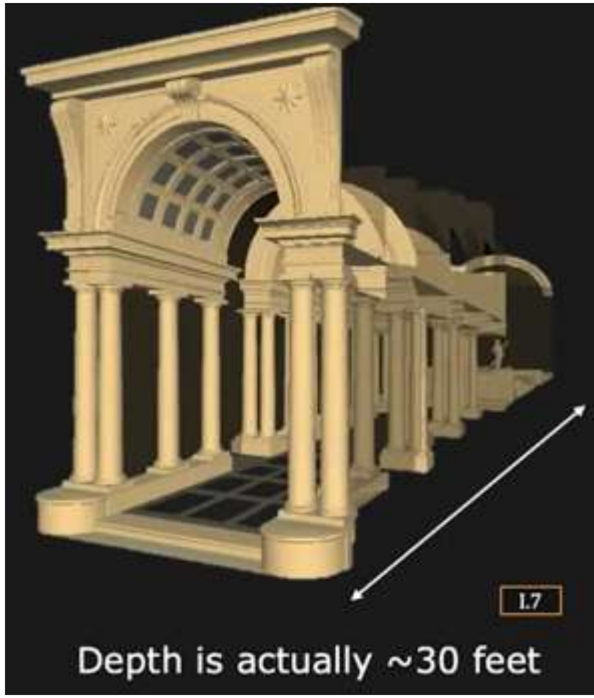
Depth appears to be ~155 feet



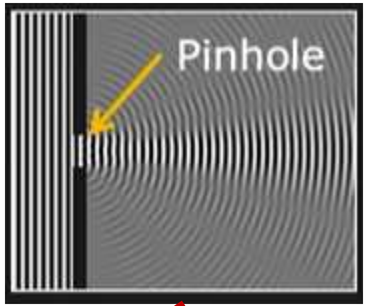
Depth is actually ~30 feet

False Perspective

- The reason of this effect or illusion is because the pillars of the archway are actually getting smaller with distance away from the observer.
- It is actually a tapered archway.
- This leads you to believe that the object is much further away than it actually is.



What is the Ideal Pinhole Size?



- The pinhole must be tiny but if it is too tiny it will cause diffraction.
- If there is an opening from which light will pass, there is going to be a bending of these light waves at the edge of the periphery of the opening.
- The smaller the opening gets; there will be more effect of the bending than the light actually passing through.

What is the Ideal Pinhole Size?

- Ideal Pinhole Diameter:

$$d \approx 2\sqrt{f\lambda}$$

f: effective focal length

λ: wavelength of light

- So, if we are working with visible light images, we can use the wavelength that ranges from 400 to 700 nanometer

What about Exposure Time?

- Flatiron Building in New York.
- The image is focused almost everywhere.
- So well-designed pinhole cameras tend to create focused-everywhere images.

$$f = 73\,mm \text{ and } d = 0.2\,mm$$

$$\text{Exposure } T = 12\,s$$

- Since Pinhole pass less light and hence require long exposures to capture bright images.

Need for lenses

- Since pinhole captures very little light, it lets very little light through.
- As a result of that the exposure times tend to be much longer.
- So that the detector that you are using to capture the image has enough photons that arrive on it.

- So, for any real application of computer vision, waiting for 12 seconds to capture a single frame is not going to work.
- That is why we use lenses.

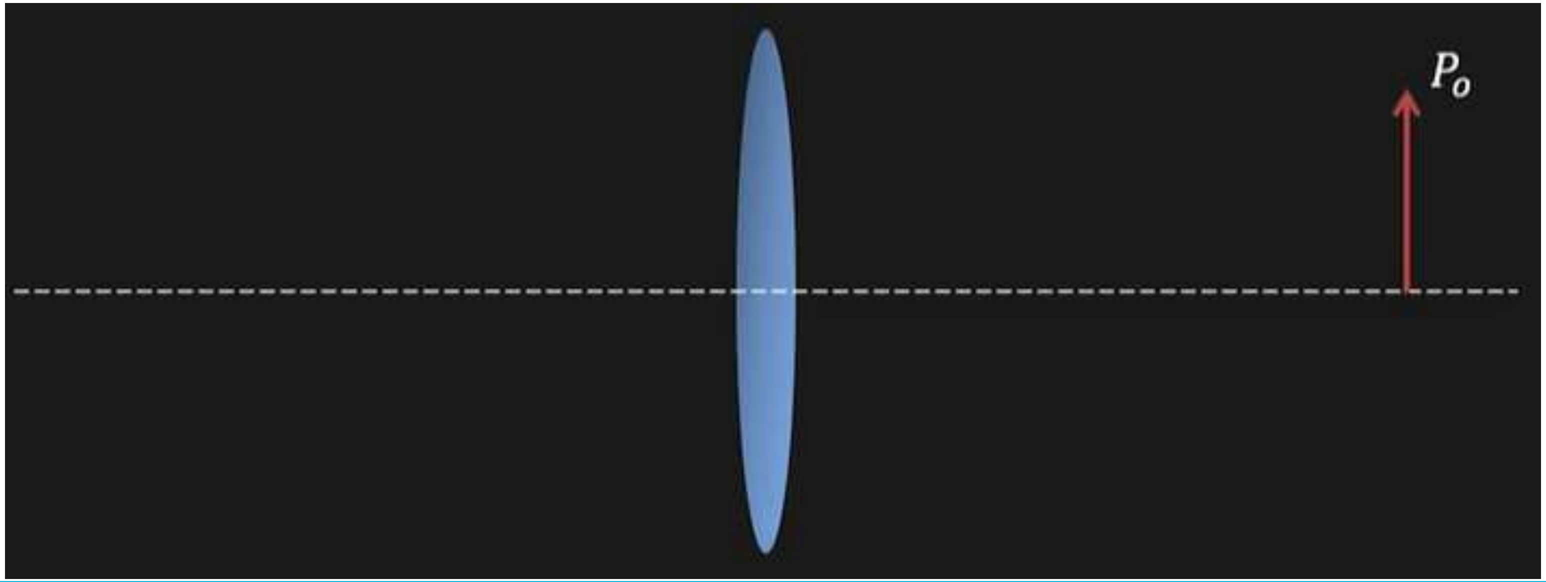
Image Formation Using Lenses

Lenses

- Now we will see how an image is formed using a lens.
- The lens performs the same projection as a pinhole which is perspective projection but gathers a lot more light

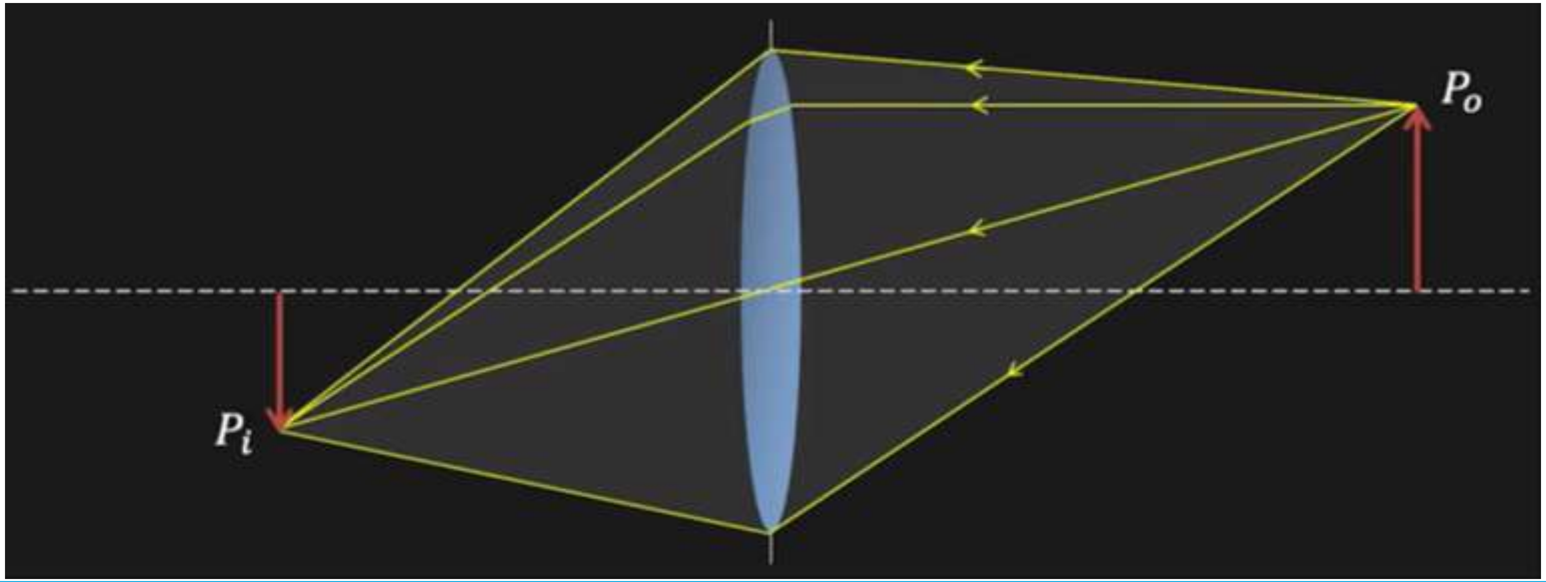
Lenses

- Consider point P_0 in the scene



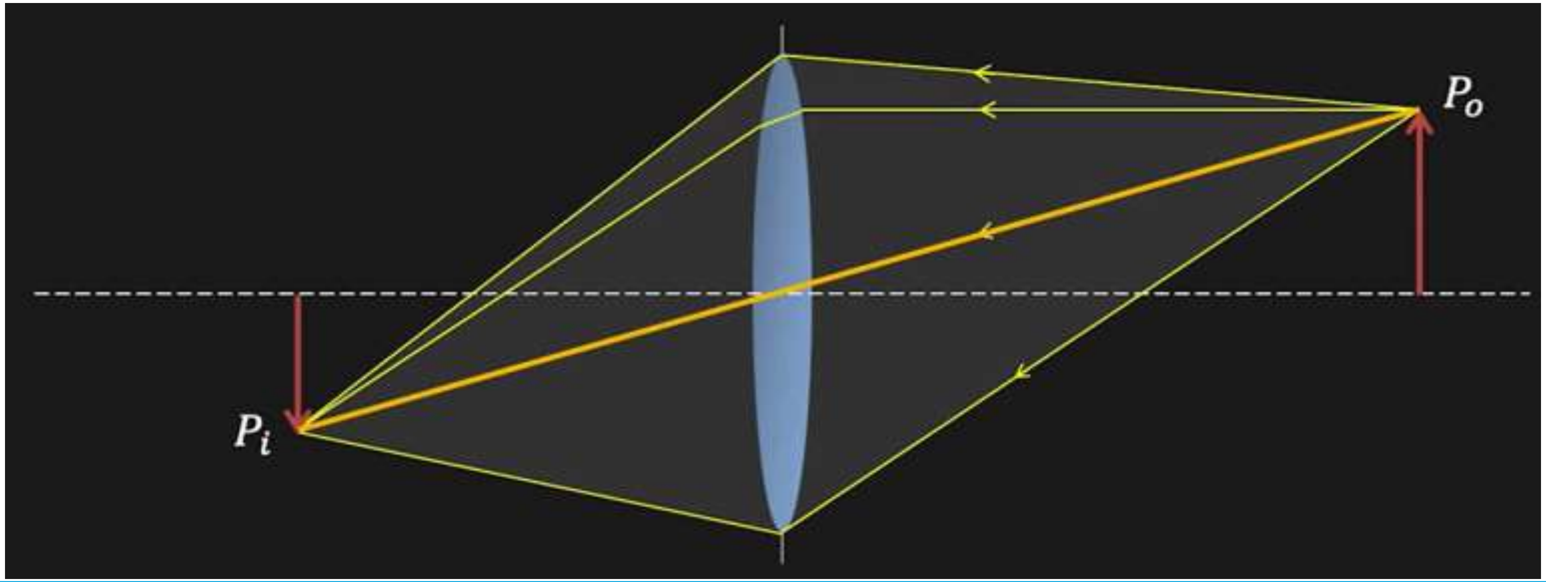
Lenses

- All the rays of light received by the lens from the point P_0 are refracted or bent by the lens to converge at the point P_i
- P_i is the point where P_0 is going to be focused behind the lens.
- The bending power of the lens is defined by its focal length.



Lenses

- If you compare these with a pinhole camera, the only ray of light that passes through is the orange ray.
- So, the perspective projection model remains the same except that the lens is able to gather significantly greater amount of light.



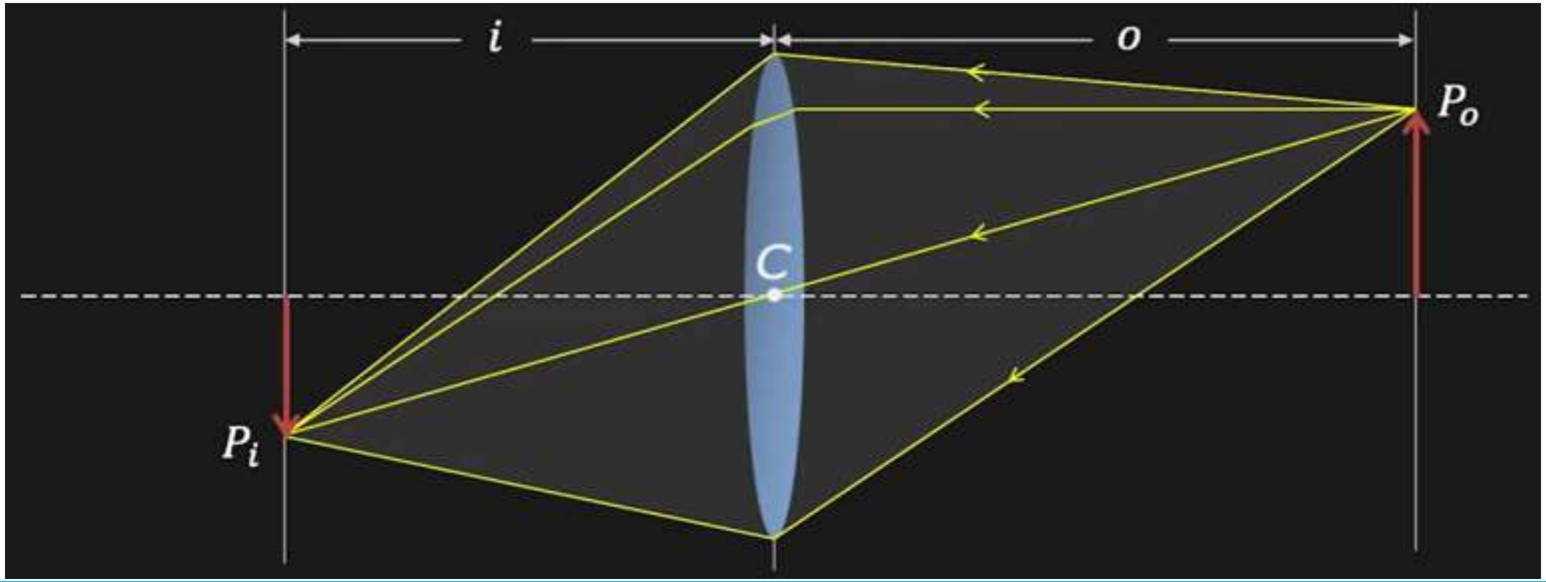
Lenses

- Let's look at the relationship between the position of the point P_0 and the position of its image P_i
- For example, if lens has $f = 50\, mm$ and distance of object from the lens is $o = 300\, mm$ then $i = 600\, mm$

f: focal length

i: image distance

o: object distance



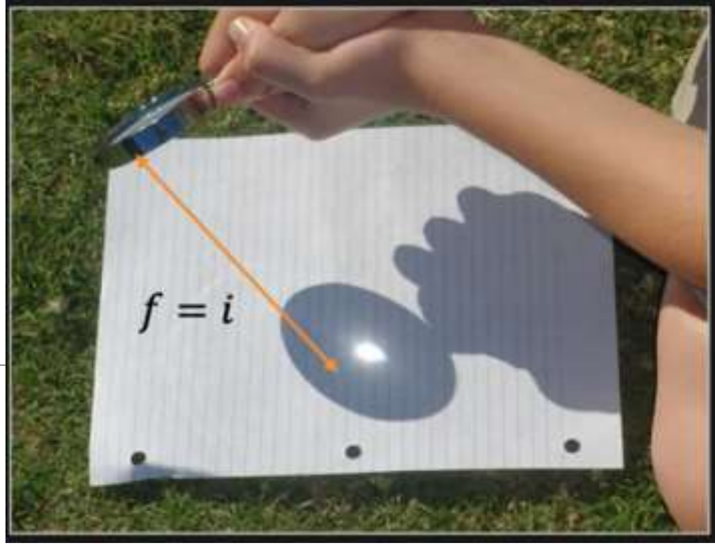
$$\frac{1}{i} + \frac{1}{o} = \frac{1}{f}$$

How to Find the Focal Length?

$$\frac{1}{i}+\frac{1}{o}=\frac{1}{f} \Longrightarrow if\ o=\infty, then\ f=i$$

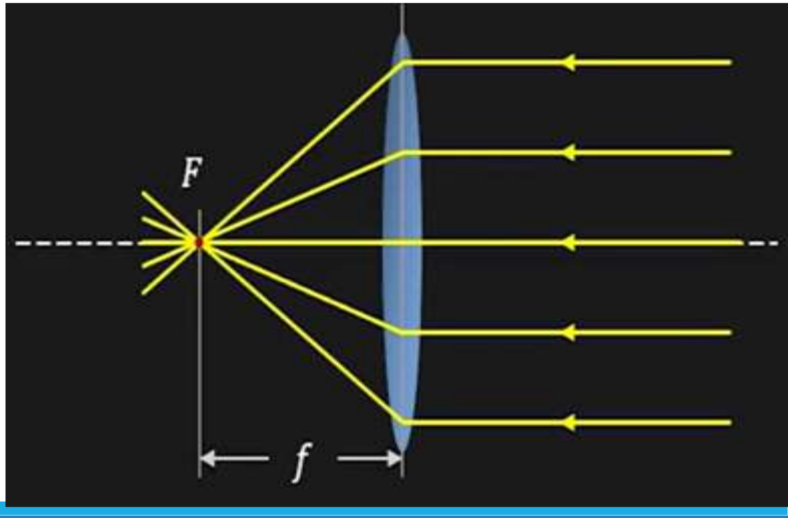
- According to Gaussian Lens Law, if the distance of the object is very far away from the lens, then the image is formed at the focal length.

How to Find the Focal Length?



- By showing the lens an object that is really far away such as the sun, a very distant point light source and looking at where the image of that point source is formed on a sheet of paper, the distance between the focused image and the lens is the focal length.

Focal Length: Distance at which incoming rays that are parallel to the optical axis converge.



How to Find the Focal Length?

- The bending power or focal length of the lens is determined by few factors.
    - The material that the lens is made of such as plastic or glass.
        - So, the refractive index of the material determines f to a great extent.
    - The shape of the lens that has two curved surfaces.
        - The radii of curvature of these two surfaces determine the focal length of the lens

Thank you