Computer Vision (CSE3010)

Dr. Susant Kumar Panigrahi
Assistant Professor
School of Electrical & Electronics Engineering

Module-1 Syllabus

Digital Image Formation And Low Level Processing:

- Overview and State-of-the-art, Fundamentals of Image Formation, Transformation: Orthogonal, Euclidean, Affine, Projective, Fourier Transform,
- Convolution and Filtering, Image Enhancement, Restoration, Histogram Processing.

Module-2 Syllabus

Depth Estimation And Multi-Camera Views:

Depth Estimation and Multi-Camera Views: Perspective, Binocular Stereopsis: Camera and Epipolar Geometry; Homography, Rectification, DLT, RANSAC, 3-D reconstruction framework; Auto-calibration. apparel.

Module-3 Syllabus

Feature Extraction And Image Segmentation:

- **Feature Extraction**: Edges - Canny, LOG, DOG; Line detectors (Hough Transform), Corners - Harris and Hessian Affine, Orientation Histogram, SIFT, SURF, HOG, GLOH, Scale-Space Analysis- Image Pyramids and Gaussian derivative filters, Gabor Filters and DWT.
- **Image Segmentation:** Region Growing, Edge Based approaches to segmentation, Graph-Cut, Mean-Shift, MRFs, Texture Segmentation; Object detection.

Module-4 Syllabus

Pattern Analysis And Motion Analysis:

- **Pattern Analysis**: Clustering: K-Means, K-Medoids, Mixture of Gaussians, Classification: Discriminant Function, Supervised, Un-supervised, Semi-supervised; Classifiers: Bayes, KNN, ANN models;
- **Dimensionality Reduction**: PCA, LDA, ICA; Non-parametric methods. Motion Analysis: Background Subtraction and Modelling, Optical Flow, KLT, Spatio-Temporal Analysis, Dynamic Stereo; Motion parameter estimation.

Module-5 Syllabus

Shape From X:

Light at Surfaces; Phong Model; Reflectance Map;

Albedo estimation; Photometric Stereo; Use of Surface Smoothness

Constraint; Shape from Texture, color, motion and edges.

Guest Lecture on Contemporary Topics

Text Books

1. Richard Szeliski, Computer Vision: Algorithms and Applications, Springer-Verlag London Limited 2011.
2. Computer Vision: A Modern Approach, D. A. Forsyth, J. Ponce, Pearson Education, 2003.

Reference Book(s):

1. R.C. Gonzalez and R.E. Woods, Digital Image Processing, Addison- Wesley, 1992.
2. Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision, Second Edition, Cambridge University Press, March 2004.
3. K. Fukunaga; Introduction to Statistical Pattern Recognition, Second Edition, Academic Press, Morgan Kaufmann, 1990.

Required Tools/Software/IDLE:

1. Python/jupyter-notebook/google-colab
2. OpenCV
3. MATLAB

Indicative List of Experiments:

1. Implement image preprocessing and Edge
2. Implement camera calibration methods
3. Implement Projection
4. Determine depth map from Stereo pair
5. Construct 3D model from Stereo pair
6. Implement Segmentation methods
7. Construct 3D model from defocus image
8. Construct 3D model from Images
9. Implement optical flow method
10. Implement object detection and tracking from video
11. Face detection and Recognition
12. Object detection from dynamic Background for Surveillance
13. Content based video retrieval
14. Construct 3D model from single image

Computer Vision

Unit – 02

Depth Estimation And Multi-Camera Views (Linear Camera Model and Camera Calibration)

Standing on the shoulder of Giants: Ref: Few Slides borrowed from:

1. Prof. Shree Nayar, *First Principles of Computer Vision is a lecture series*.
2. Prof. Mubarak Shah, *Computer Vision Video Lectures*.

Computer vision is the enterprise of building machines that "See"

Cameras Maps 3D objects to 2D (Loss of dimensionality): No Depth information.

Analyzing 3D world from 2D information is always difficult. Thus computer vision is challenging.

Camera Model and Camera Calibration

Determine intrinsic and extrinsic parameters of camera

✓ Extrinsic Parameters:

- 3D Location and Orientation of camera.

✓ Intrinsic Parameters:

- Focal Length
- The size of the pixel.

Applications: Object Transfer

Source Image

Target Image

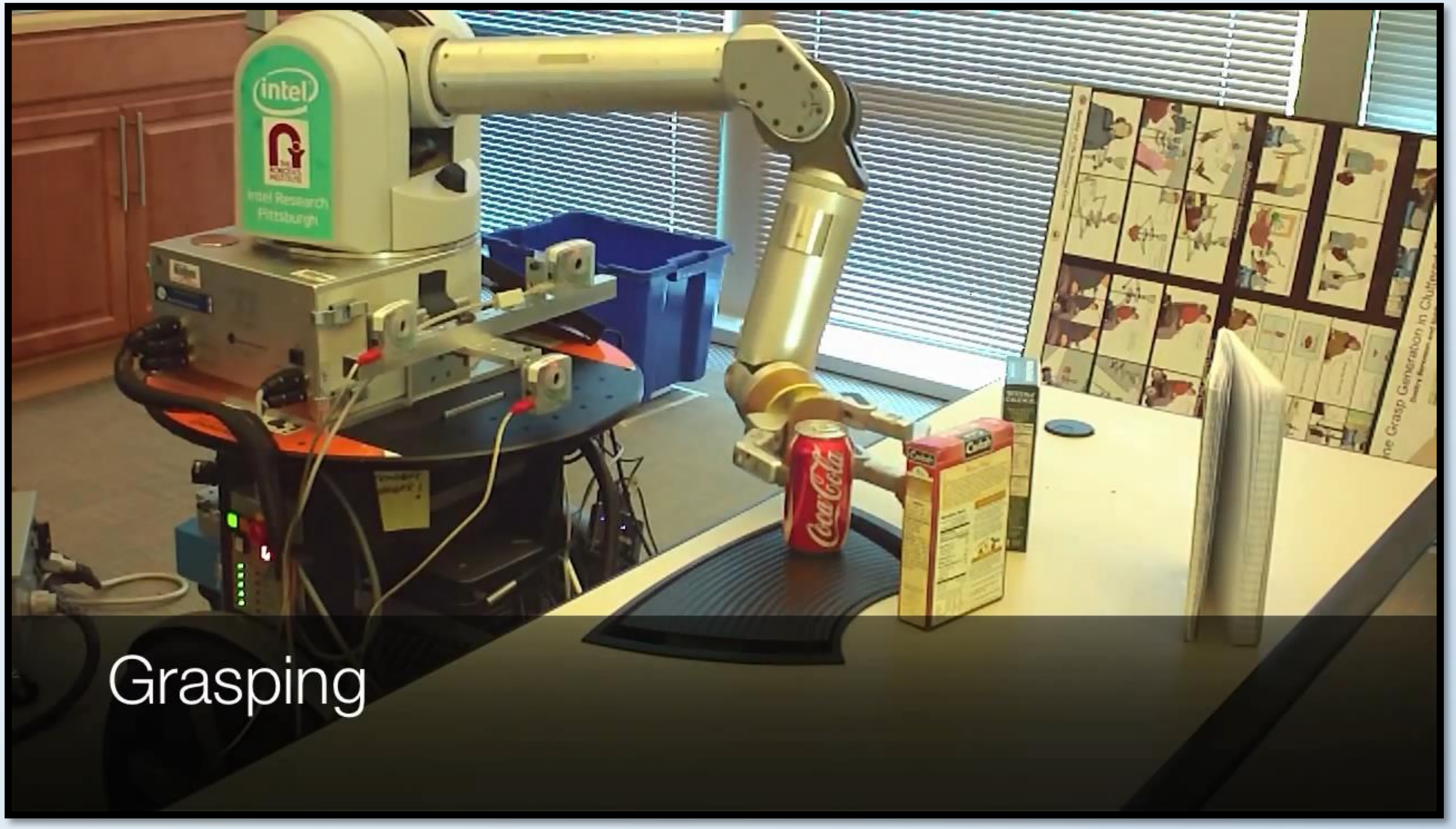
Applications: Object Transfer

Source Image

Target Image

Applications: Pose Estimation

Given 3D model of object, and its image (2D projection) determine the location and orientation (translation & rotation) of object such that when projected on the image plane it will match with the image.



Source: http://www.youtube.com/watch?v=ZNHRH00UMvk

Image Formation

Image: Projection of 3D scene onto 2D plane. We need to understand the geometric and photometric relationship between the scene and its image.

Given a point in the scene; we need to know where it ends up in the image plane.

Given the brightness or appearance of the scene; we need to find the brightness or appearance of the resultant image.

Require: Camera Model and Camera Calibration

Before going for camera modeling it is better to have an overview of image formation in a camera based on its type.

1. Image formation in Pinhole Camera
2. Prospective Projection
3. Image Formation using Lenses.

Image Formation in Pinhole camera

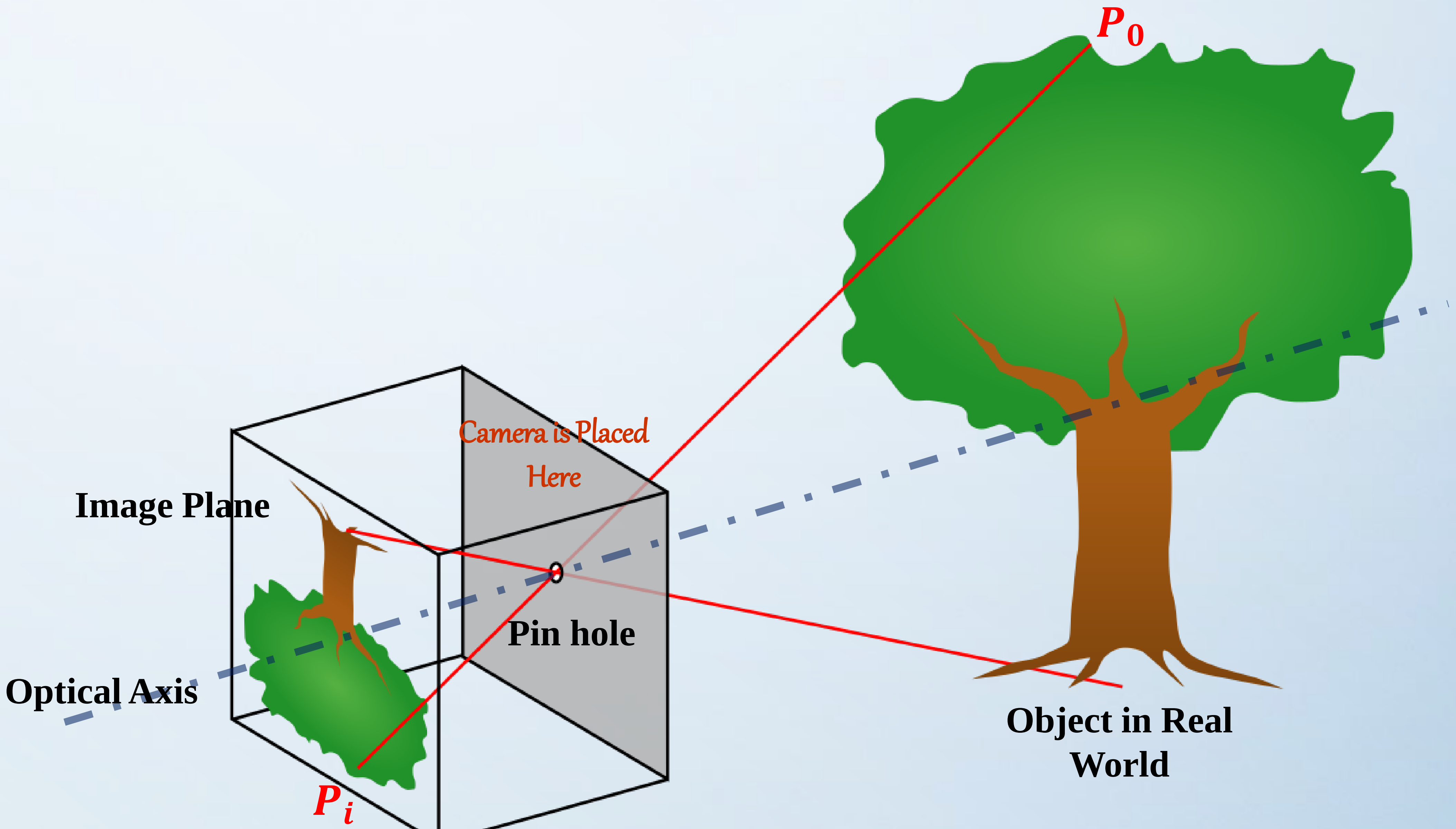
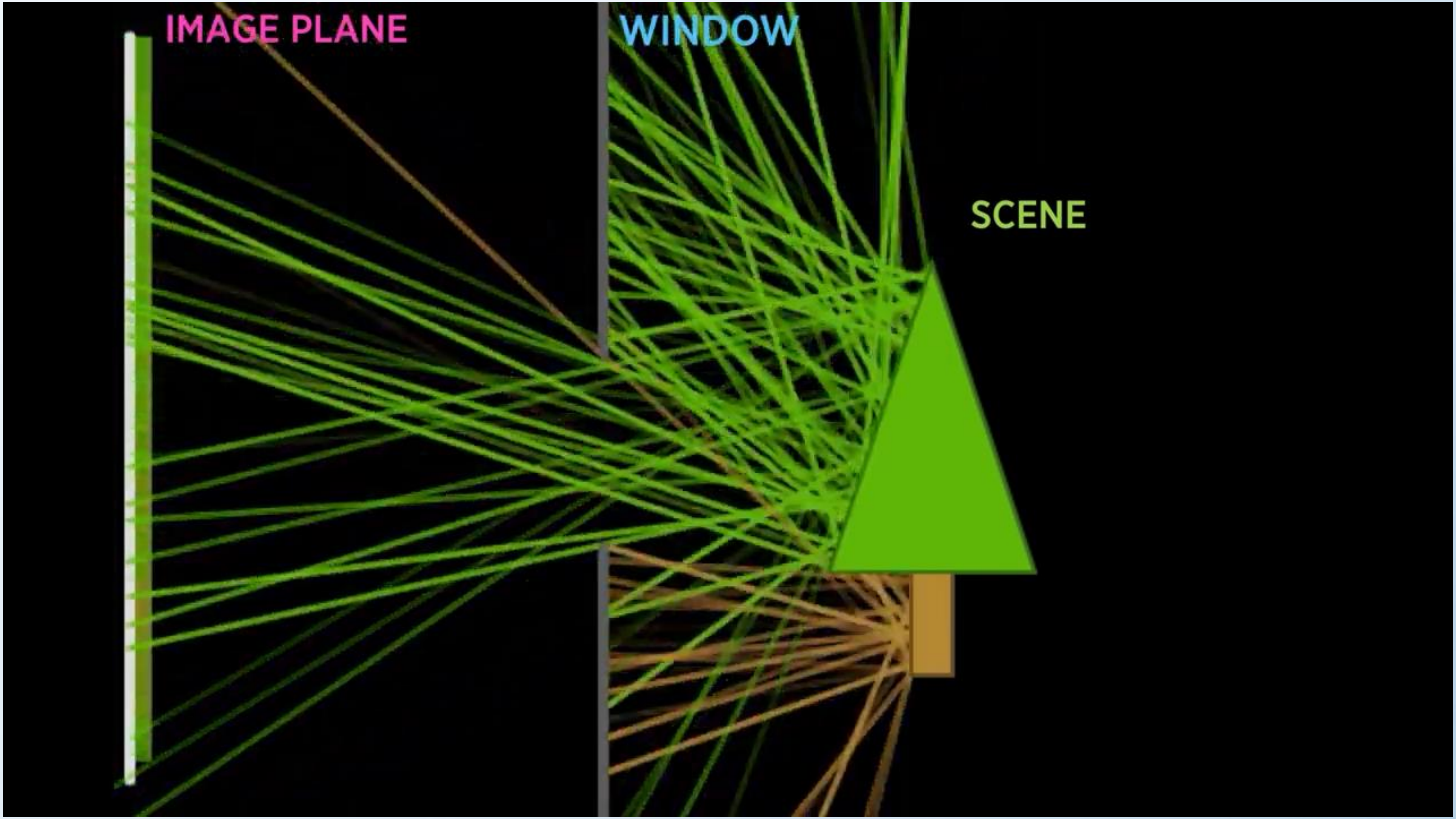
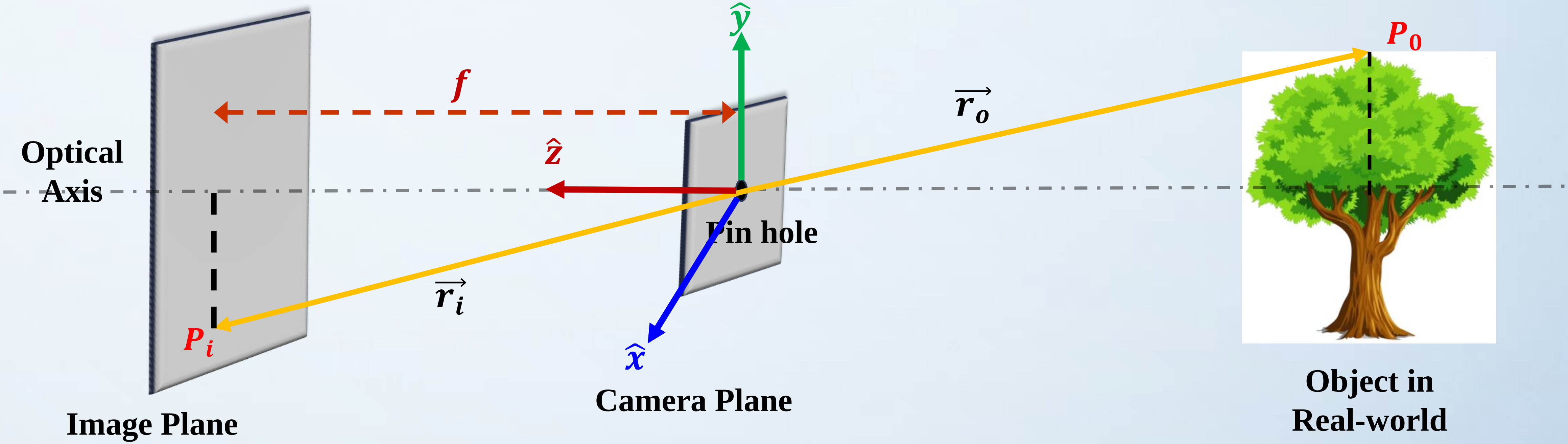


Image Formation in Pinhole camera



Perspective Projection and Pinhole Camera



f : Effective Focal Length:

$$\vec{r_o} = (x_o, y_o, z_o)$$

$$\vec{r_i} = (x_i, y_i, f)$$

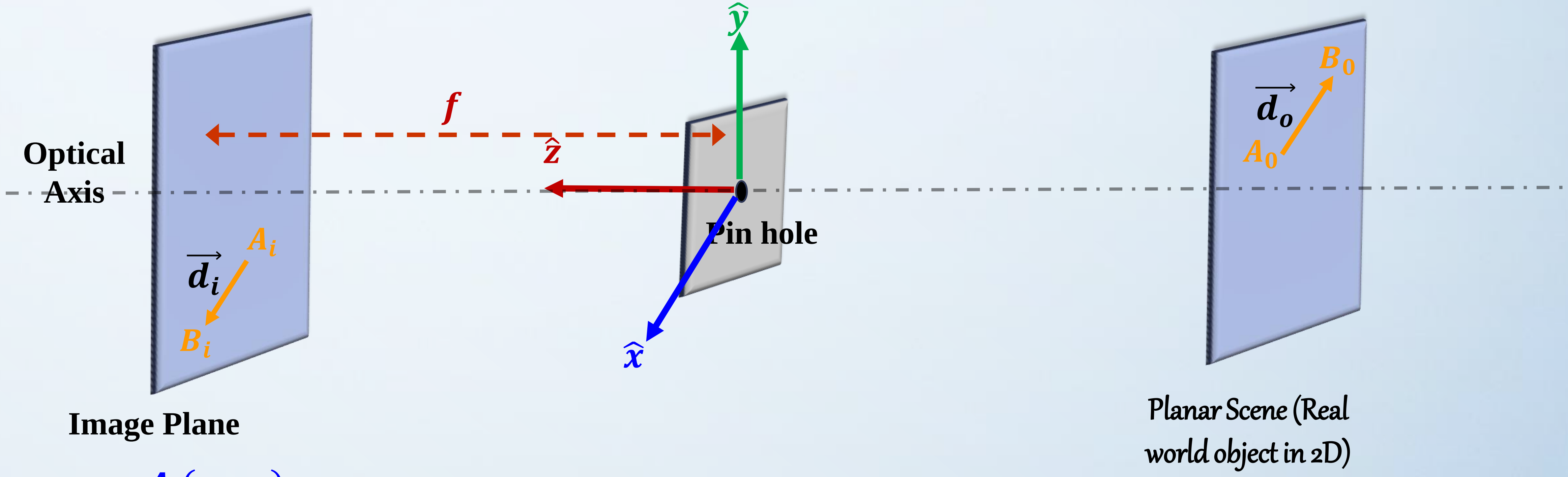
Using triangle similarity:

$$\frac{\vec{r_i}}{f} = \frac{\vec{r_o}}{z_o}$$

$$\frac{x_i}{f} = \frac{x_o}{z_o}, \quad \frac{y_i}{f} = \frac{y_o}{z_o}$$

Irrespective of the object in the scene the Z-component of image P_i will always be the effective focal length.

Image Magnification



$A_i(x_i, y_i)$

$B_i(x_i + \delta x_i, y_i + \delta y_i)$

$A_0(x_o, y_o, z_o)$

$B_0(x_o + \delta x_o, y_o + \delta y_o, z_o + \delta z_o)$

Magnification: $|m| = \frac{\|\vec{d_i}\|}{\|\vec{d_o}\|} = \frac{\sqrt{\delta x_i^2 + \delta y_i^2}}{\sqrt{\delta x_o^2 + \delta y_o^2}}$

Image Magnification

From perspective projection:

$$\frac{x_i}{f} = \frac{x_o}{z_o} \;,\quad \frac{y_i}{f} = \frac{y_o}{z_o} \quad \text{.....................} \; (A)$$

$$\frac{x_i+\delta x_i}{f} = \frac{x_o+\delta x_o}{z_o} \;,\quad \frac{y_i+\delta y_i}{f} = \frac{y_o+\delta y_o}{z_o} \quad \text{.....................} \; (B)$$

From (A) and (B):

$$\frac{\delta x_i}{f} = \frac{\delta x_o}{z_o} \;,\quad \frac{\delta y_i}{f} = \frac{\delta y_o}{z_o}$$

Magnification: $|\boldsymbol{m}| = \frac{\|\vec{\boldsymbol{d_i}}\|}{\|\vec{\boldsymbol{d_o}}\|} = \frac{\sqrt{\delta x_i^2+\delta y_i^2}}{\sqrt{\delta x_o^2+\delta y_o^2}} = \left|\frac{f}{z_o}\right|$

$\boldsymbol{m} = \frac{f}{z_0}$; m is negative when image is inverted.

Image Magnification

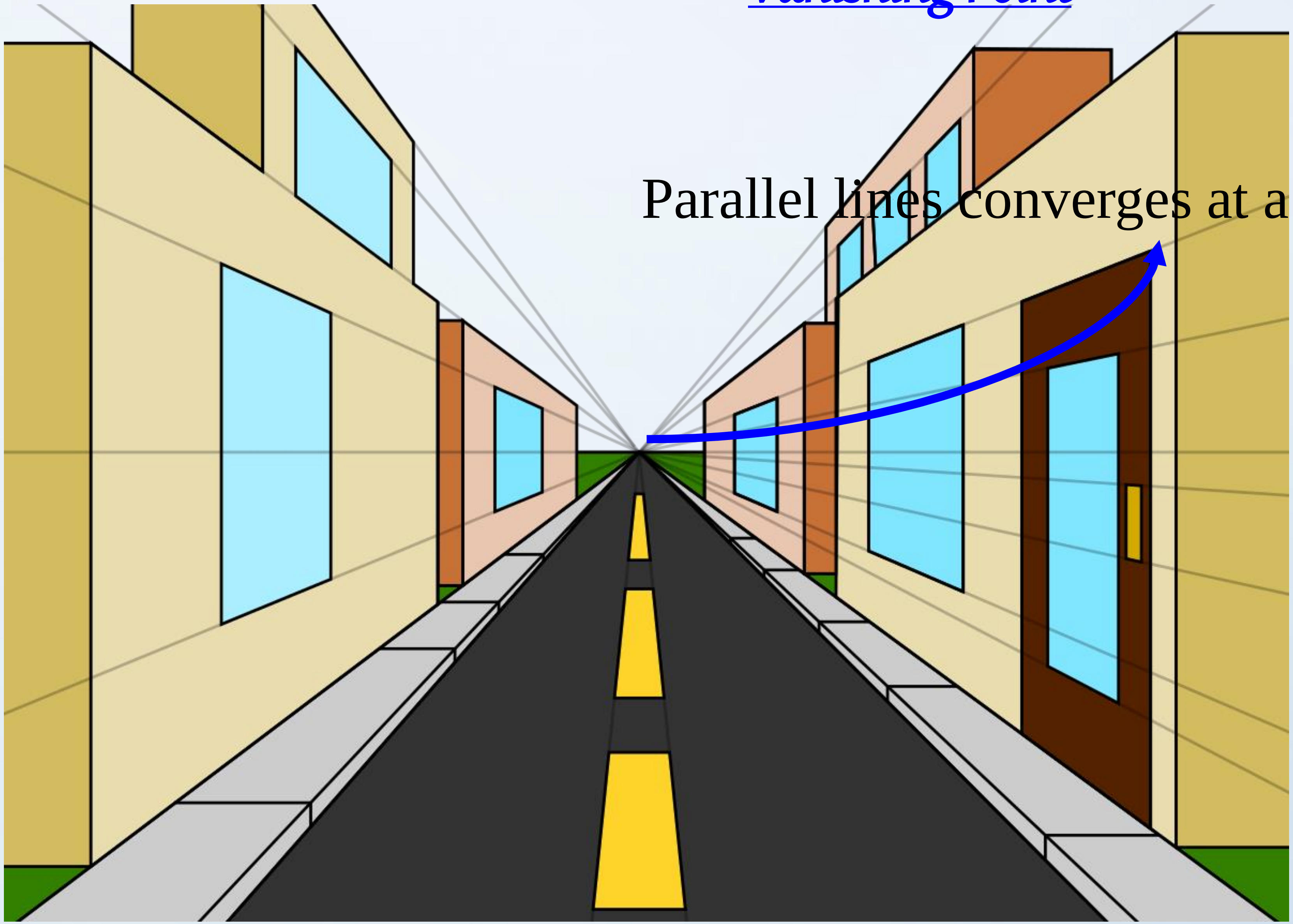
$m = \frac{f}{z_0}$; Image Size is inversely propositional to depth.

Remarks on Image Magnification

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

- ✓ A small object at a distance greater than the depth of the object will magnify equally as the scene on to the image.

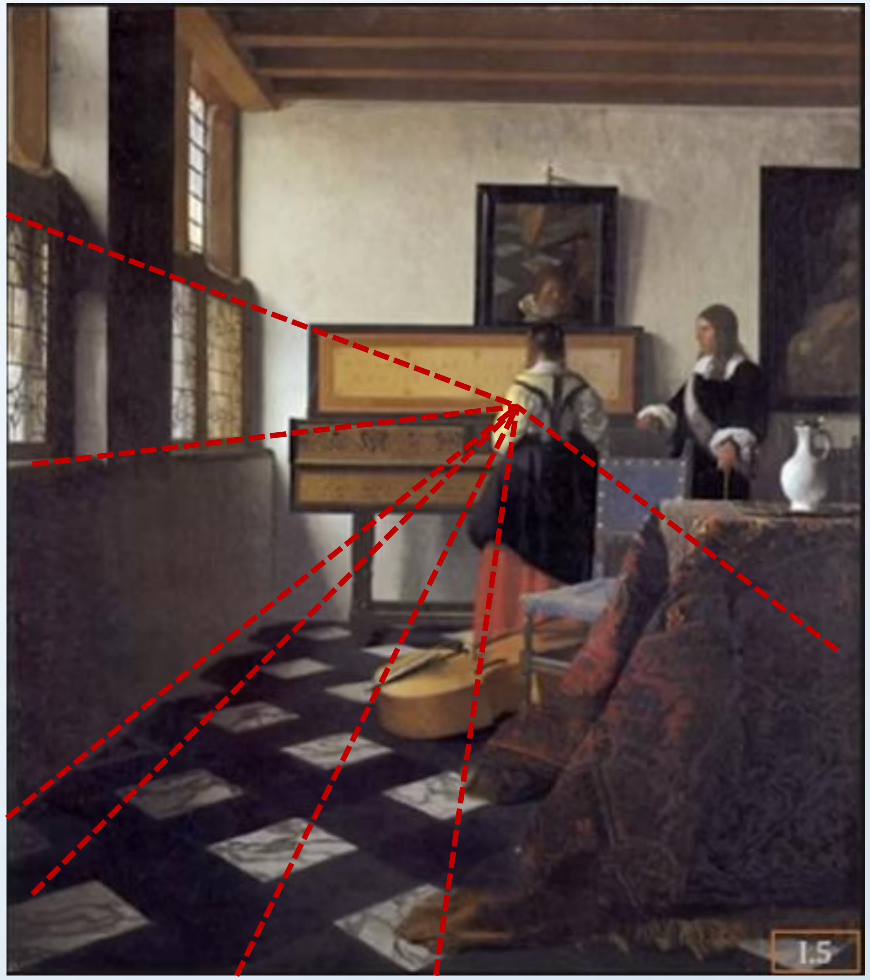
Vanishing Point



Parallel lines converges at a single point.

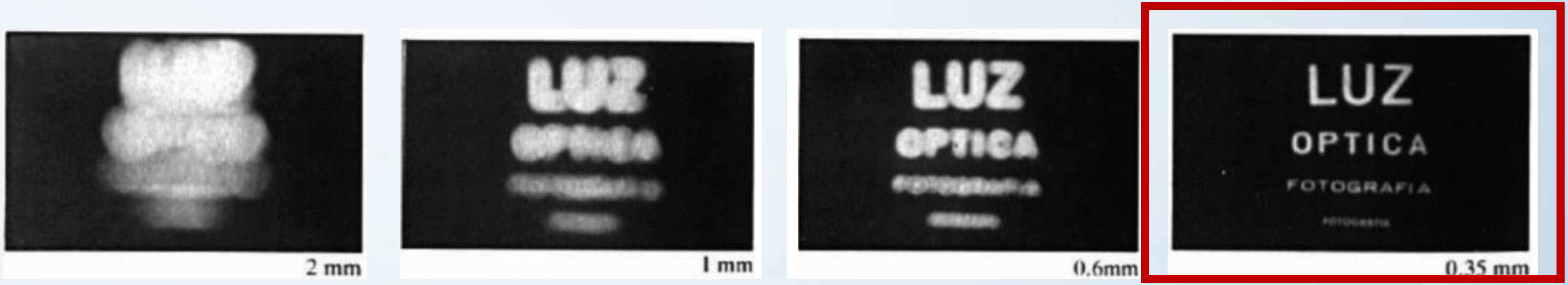
Location of the **vanishing point** depends on the **orientation** of parallel straight line.

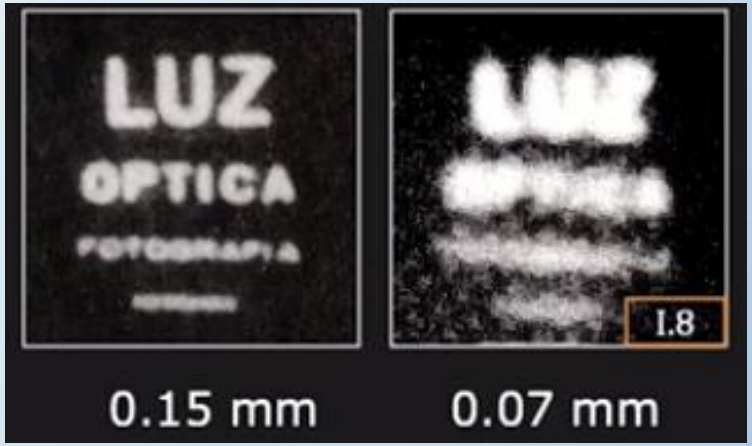
Use of Vanishing Point in Art



The Music Lesion: Johannes Vermeer, 1662-1664

What is the Ideal Pinhole Size





The Pin hole must be **tiny**
but if it is **too tiny**, then it will cause **diffraction**.

Ideal pin hole diameter:

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

Where, f: Effective Focal Length.
λ: Wavelength of Light

Long Exposure Time and Pinhole Camera

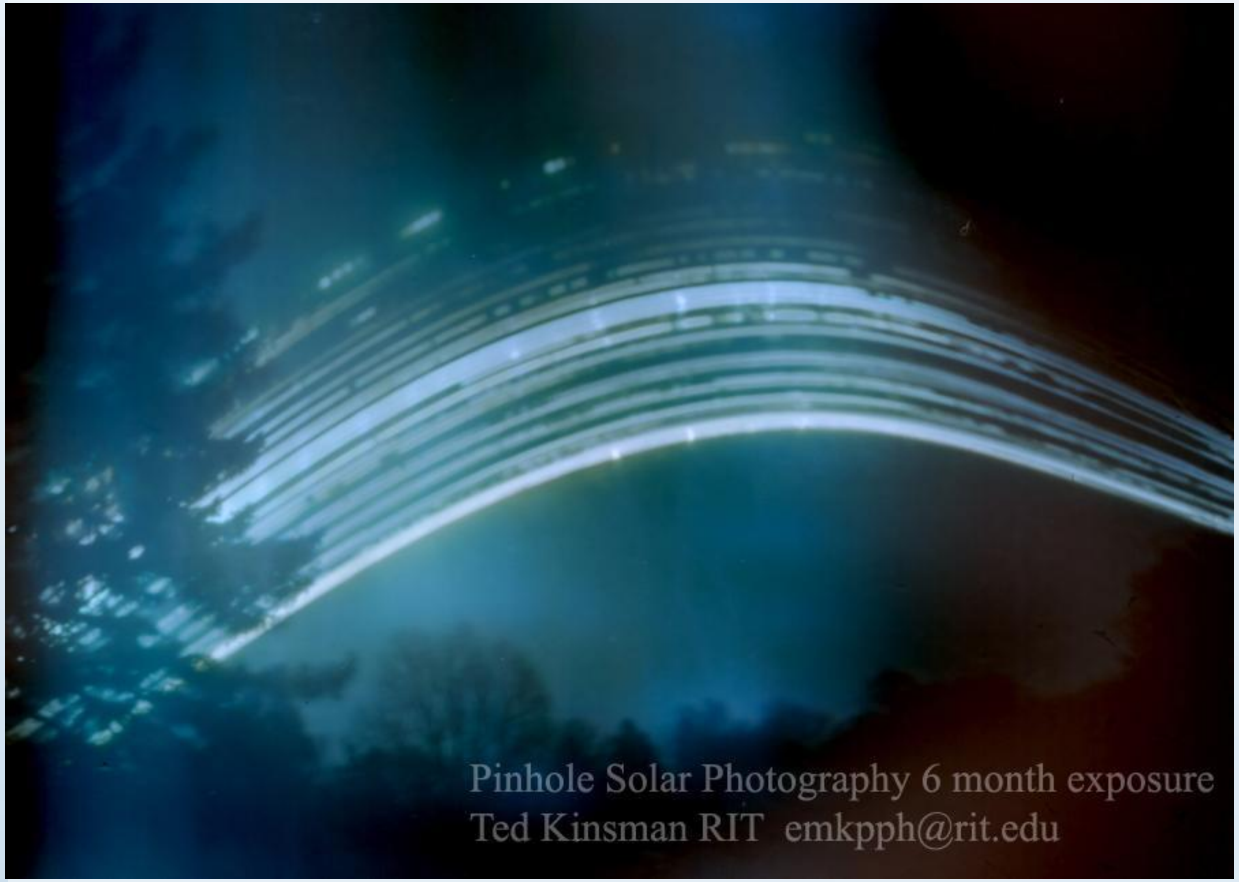
A **small** opening in Pinhole passes **less light**. Image Looks darker and distorted. Hence require **long exposures** to capture bright images.

Long exposure adds **motion blur** and other distortion, **impractical** for real-time computer vision task.

$f = 73mm, d = 0.2mm$

Exposure Time $\boldsymbol{T = 12sec}$

Long Exposure Time and Pinhole Camera



6 month exposure with a pinhole camera, showing the path of the Sun during this time.

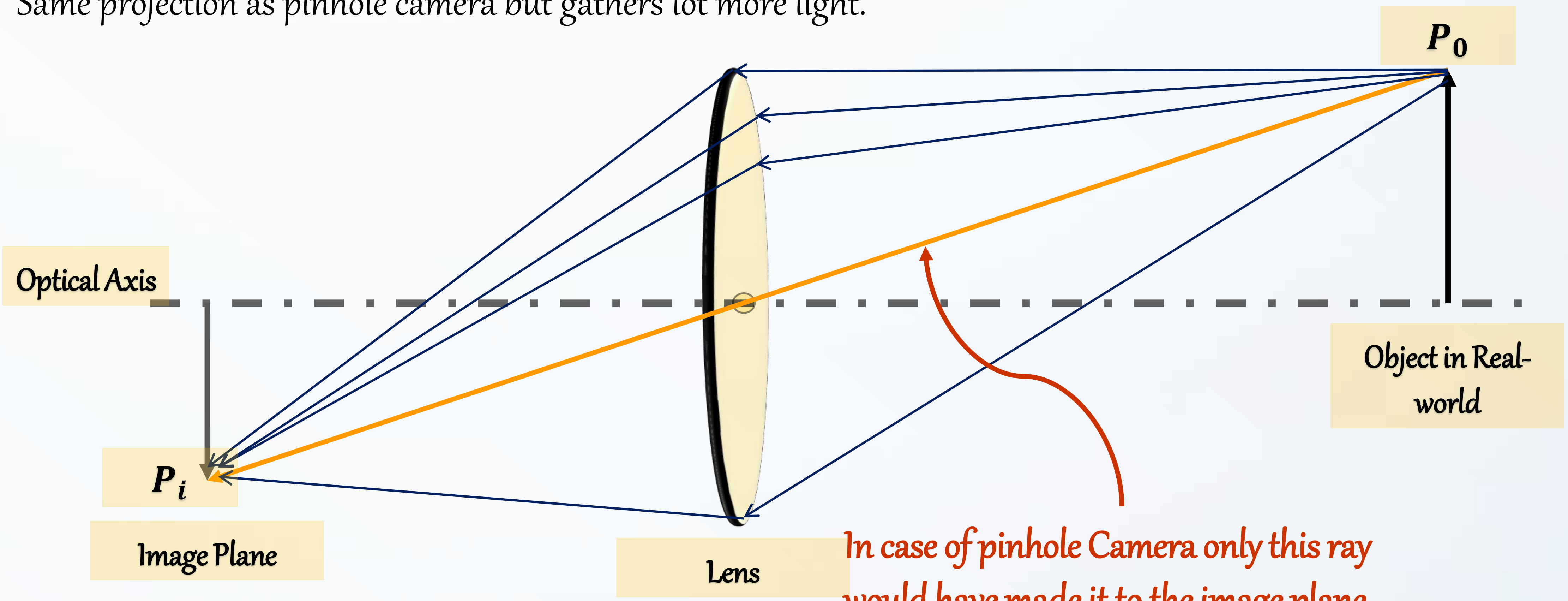
Longest Pinhole Exposure Ever Is Also Most Beautiful

This longest-ever exposure shows 34months of New York

Source: https://www.pinterest.com/pin/145663987648071651/

Image Formation in Lenses

✓ Same projection as pinhole camera but gathers lot more light.



In case of pinhole Camera only this ray would have made it to the image plane for a particular point in the world coordinate.