Machine Vision & Perception

Formation of Images

Mirza Farhan Bin Tarek [1]

[1]Department of Computer Science and Engineering
BRAC University

CSE461: Introduction to Robotics

Table of Contents

Table of Contents

Instruments of Vision

(a) Robot with Camera

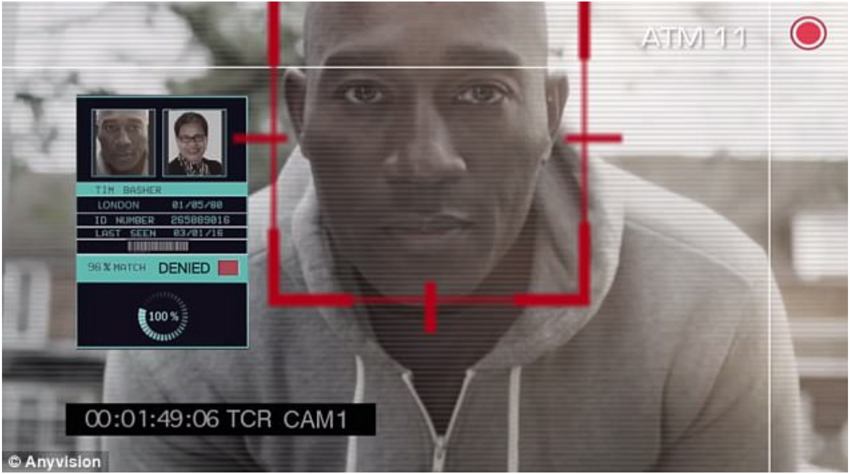
(b) DSLR Camera

Figure: Different Instruments of Capturing Image

Applications of Computer Vision

TONS OF USES OF COMPUTER VISION!!!

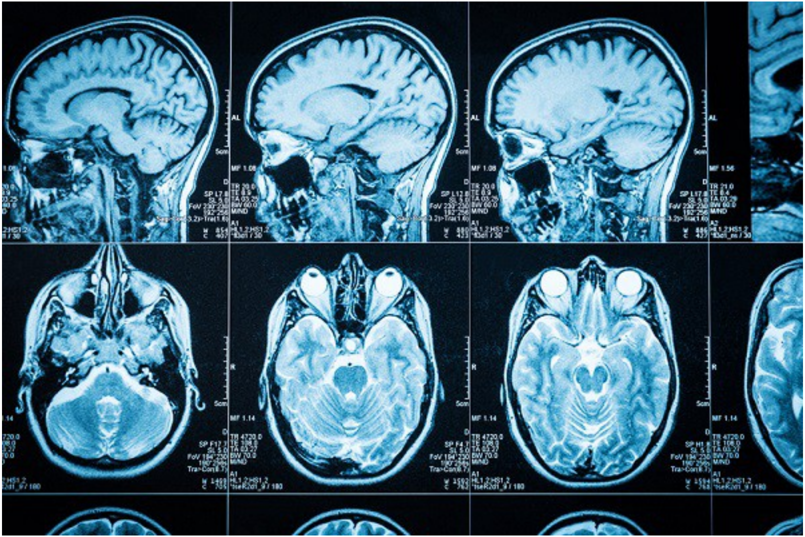
Face Recognition



Motion Capturing

Pose Estimation

Medical Image Processing



What we will learn

- Basic geometric models of how images are formed
- Fundamentals of image sensors
- Fundamentals of photography devices and imaging techniques

Table of Contents

Perspective Projection

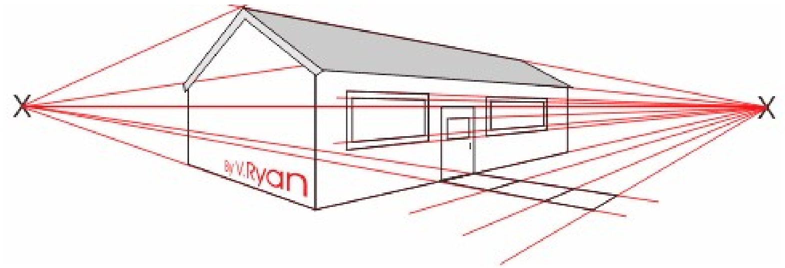
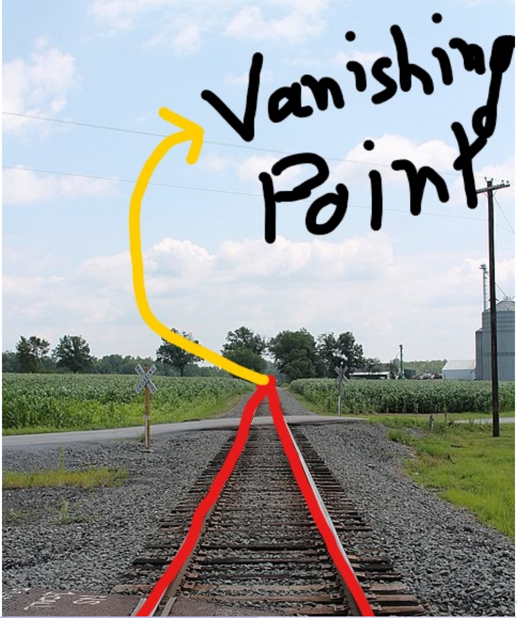


Figure: Notice the Geometry

Perspective Projection

- Straight lines stay straight
- However, parallel lines does not remain parallel


Vanishing
Point

Vanishing point

- Parallel lines intersect at the *Vanishing Point*
- In our regular geometry, we say 'parallel lines intersect at infinity'
- The vanishing point is THAT point
- Each direction has exactly one vanishing point

Motivation for Projective Geometry

- Euclidean geometry is not optimal in explaining these phenomena including central projection. The math can get difficult.
- That is why, we can extend the Euclidean space by adding points at infinity and say, the parallel lines meet at these points. This is how we *extend* the Euclidean Space to *Projective Space*
- **Projective geometry** is an alternative representation of geometric objects and transformations.

Homogeneous Coordinates

Definition

The representation x of a geometric object (e.g. a point) is **homogeneous** if x and λx represent the same object for $\lambda \neq 0$. For example in Euclidean form, $x \neq \lambda x$ whereas in H.C $x = \lambda x$

[0] We will abbreviate Homogeneous Coordinates as H.C. in the next few slides

Homogeneous Coordinates

- H.C. uses (n+1) dimensions to represent an n-dimension Euclidean point
- The dimension (n+1) is set to 1
- If an Euclidean Point is $x = \begin{bmatrix} x \\ y \end{bmatrix}$ then it's H.C. would be $x = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

Homogeneous Coordinates

Conversion:

- heterogeneous → homogeneous

$$\begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- homogeneous → heterogeneous

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow \begin{bmatrix} x/w \\ y/w \end{bmatrix}$$

- scale invariance

$$\begin{bmatrix} x & y & w \end{bmatrix}^\top = \lambda \begin{bmatrix} x & y & w \end{bmatrix}^\top$$

Special points:

- point at infinity

$$\begin{bmatrix} x & y & 0 \end{bmatrix}$$

- undefined

$$\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

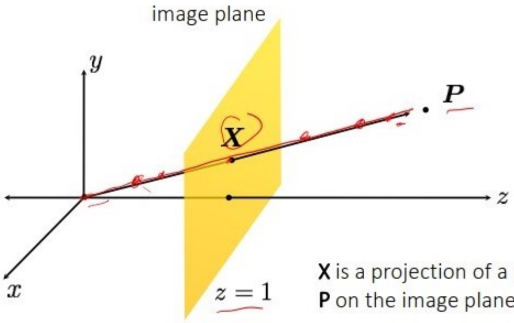
[0] Computational Photography Fall 2020, CMU

Projective Geometry

image point in pixel coordinates $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$

$$\Updownarrow$$

image point in homogeneous coordinates $\boldsymbol{X} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$



X is a projection of a point **P** on the image plane

[0] Computational Photography Fall 2020, CMU

Homogeneous Coordinates of a 3D Point

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \sim \begin{bmatrix} ax \\ ay \\ az \\ a \end{bmatrix}$$

What is Camera?

Definition of Camera

A camera is a mapping between 2D projection in image plane and the 3d source object residing in a 3D object space [1].

[1] Multiple View Geometry in Computer Vision By Richard Hartley, Andrew Zisserman, Ch-06

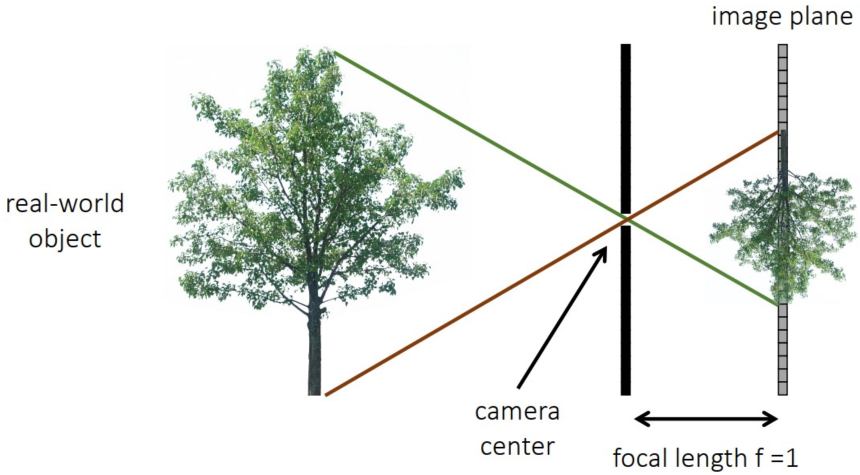
Camera as a Coordinate Transformation

$$x = PX$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Image Coordinates 3x1 | Camera Matrix 3x4 | World Coordinates 4x1

Pinhole Camera



[0] Computational Photography Fall 2020, CMU

Pinhole Camera

Definition of Pinhole Camera Model

Pinhole camera model is an ideal description of the mathematical relationship between a 3D object and its 2D projection where the camera aperture is described as an infinitesimal point and no lenses are used to focus light.[1,2].

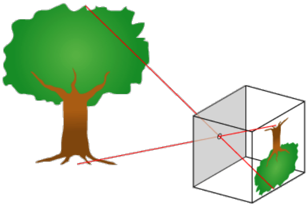


Figure: Wikipedia

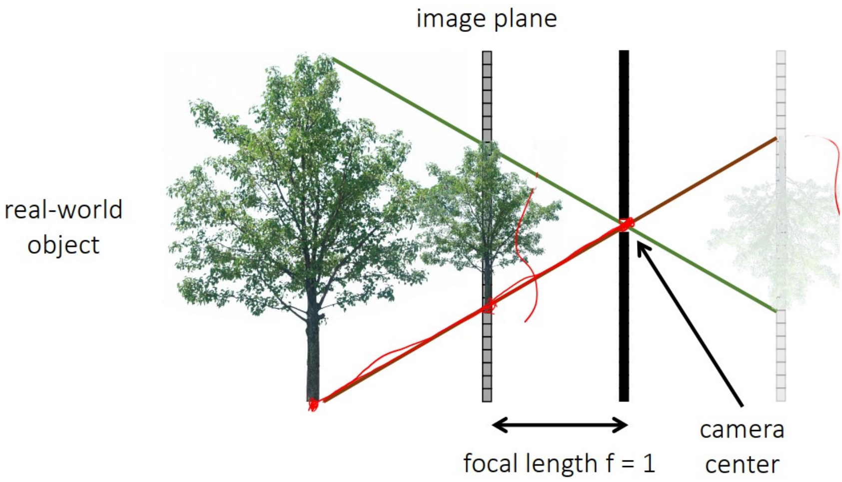
[1] Multiple View Geometry in Computer Vision By Richard Hartley, Andrew Zisserman, Ch-06

[2] Pinhole Camera Model, Wikipedia

Some Properties of the Pinhole Camera

- The pinhole is very small
- **Line-preserving**: Straight lines remain straight
- **Not length-preserving**: Size of the object is inverse proportional to the distance
- **Not angle-preserving**: Angle between lines change

Rearranged Pinhole Camera



[0]Computational Photography Fall 2020, CMU

Image Projection Using Pinhole Camera Model

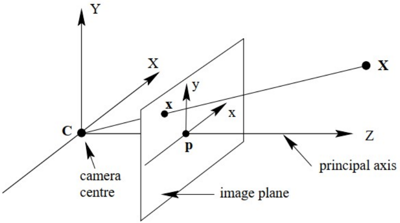


Figure: Multiple View Geometry in Computer Vision By Richard Hartley, Andrew Zisserman

- Th projection centre is called the *camera centre*
- The line from the camera centre perpendicular to the image plane is called the *principal axis*
- the point where the principal axis meets the image plane is called the *principal point*
- The plane through the camera centre parallel to the image plane is called the *principal plane of the camera*

Formulating Relationship between 2D Image and 3D Object

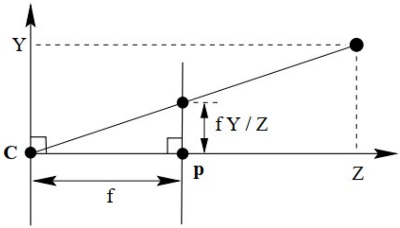


Figure: Multiple View Geometry in Computer Vision By Richard Hartley, Andrew Zisserman

Relationship

Since two triangles are similar, we can write that $[X,Y,Z]^T$ is mapped to $[X/Z,Y/Z,1]$ if $f=1$. As the image is 2D, we can ignore the final image coordinate and finally write,

$$[X \quad Y \quad Z]^T \longmapsto [X/Z \quad Y/Z]^T$$

Pinhole Camera Matrix

From the previous slide,

$$[X \quad Y \quad Z]^T \longmapsto [X/Z \quad Y/Z]^T$$

General camera model in H.C from previous slide,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

In case of the pinhole camera, the projection matrix would look like

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = [I|0]$$

Camera Matrix for Random Focal Length

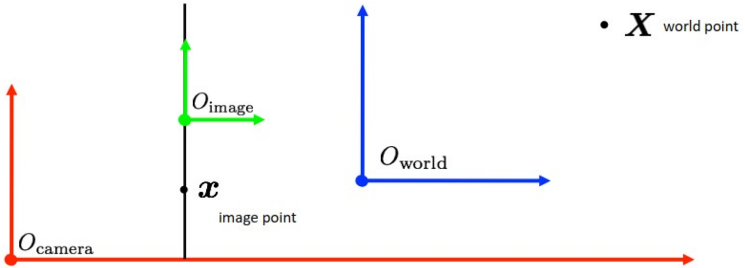
In the previous slides, we calculated the camera matrix for $f = 1$. For any focal length f, the 3D World point to 2D image point mapping would be

$$[X \quad Y \quad Z]^T \longmapsto [fX/Z \quad fY/Z]^T$$

In case of the pinhole camera, the projection matrix would look like

$$P = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Generalizing Camera Matrix



[0] Computational Photography Fall 2020, CMU

Generalizing Camera Matrix

In practical cases, the origin of image plane coordinates in image plane may not be the principal point. So in general cases we use the mapping

$$[X \quad Y \quad Z]^T \rightarrow [fX/Z + p_x \quad fY/Z + p_y]^T$$

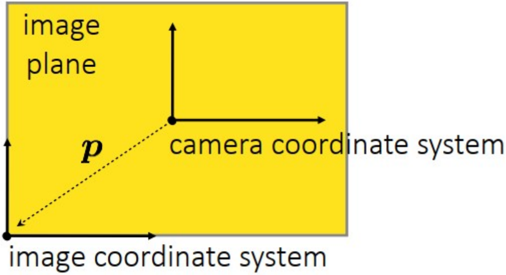


Figure: Image and Camera can sit at different coordinate systems

Generalizing Camera Matrix

We can write the new projection matrix for variable focal length as:

$$P = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

we can decompose the above matrix as

$$P = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$P = K[I|0]$$

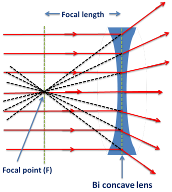
Calibration matrix,

$$K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

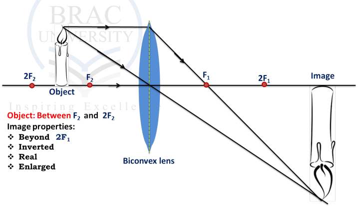
Table of Contents

Thin Lens

- Thin lens is an optical device that can converge or focus light beams into a point or diverge or disperse light beams from a point with the help of two refracting surfaces
- Two major types of lens exists: Concave and Convex



(a) Concave Lens



(b) Convex Lens

Figure: Different types of lenses

Some Terminology

- **Principal Focus** is the point where multiple parallel light beams will converge to or seem to diverge from a point. We can denote it with F. Each lens has two focal points on both sides.
- **Principal Axis** Each surface of a lens can be thought of a portion of sphere. The axis connecting the centers of these spheres is called principal axis.
- **Focal Length** is the distance between the center of the lens and either of the two principal foci. We can denote this length or distance as f.
- **Object distance** and **Image distance** We can denote the distance between the lens center and the object as O and the image as i

Lens Equation

Lens Equation

If object distance is O, image distance is i and focal length is f then,

$$\frac{1}{O} + \frac{1}{i} = \frac{1}{f} \tag{1}$$

Convex Lens

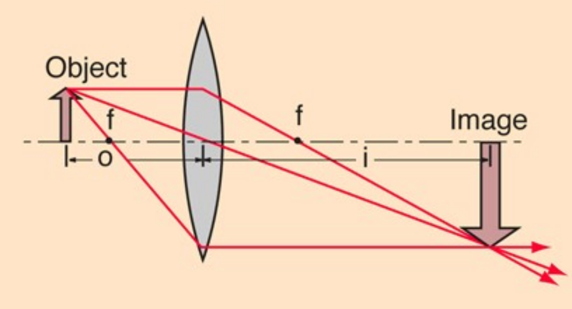


Figure: Ray diagram of a convex lens when $O > f$

Convex Lens (Continued)

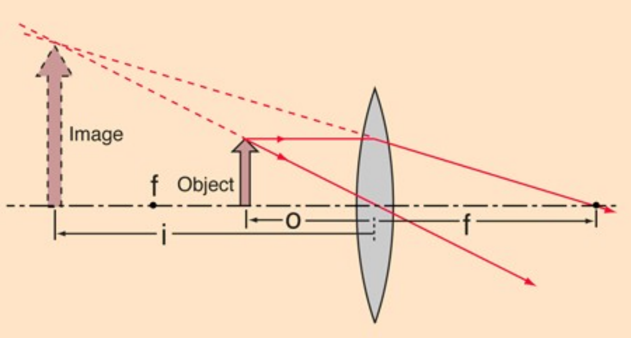


Figure: Ray diagram of a convex lens when $O < f$

Convex Lens (Continued)

Remark

When $O > f$, a **real but inverse** image forms on the opposite side of the image. If you place a screen or an image sensor there, the image will form on that plane. **THIS IS ACTUALLY A BASIC PRINCIPLE OF HOW CAMERA AND OUR EYES WORK!!**

Concave Lens

HOME TASK!!!

Section Finale

In our next topic, we will learn how vision sensors work and how a physical camera captures images.

Acknowledgement

- 15-463, 15-663, 15-862 Computational Photography Fall 2020, Lecture 19, CMU
- Multiple View Geometry in Computer Vision By Richard Hartley, Andrew Zisserman, Ch-06
- https://en.wikipedia.org/wiki/Pinhole_camera_model
- http://hyperphysics.phy-astr.gsu.edu/hbase/geoopt/raydiag.html#c2