3d computer Vision Introduction

① When all of the rays after reflecting from the object reaches the film, we get blurred image. To reduce this blurring, we use a barrier which blocks most of the rays. This barrier is called **aperture**.

② Pinhole projection model: (A) captures pencil of rays (i.e. all rays through a single point).
(B) The point is called center of projection (focal Point)
(C) The image is formed on the image plane.

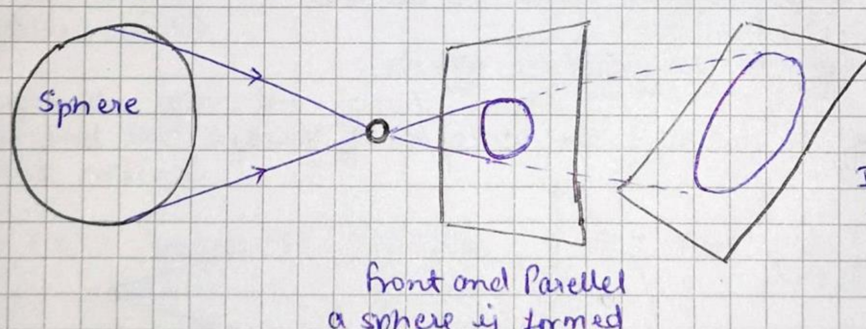
③ What we see in 3d (in real life), gets converted into a 2d image. i.e. the dimensions gets reduced (Dimensionality Reduction machine). While converting from 3d to 2d, we lose (a) angles (b) distances (lengths)

④ Projection Properties:

(A) Any point along the same ray maps to the same image point (many to one)
(B) Points → Points (But only defined for points in front of focal plane)
(C) Lines → Lines (collinearity is preserved, lines through focal points project to point)
(D) Planes → Planes (or half-planes) (But planes through focal points project to lines)

(E) Parellel lines converge at a vanishing point
  (a) Each direction in space has its own vanishing point
  (b) But parellel lines, which are parellel to the image plane remain parellel.
  (c) All directions in the same plane have vanishing points on the same line.

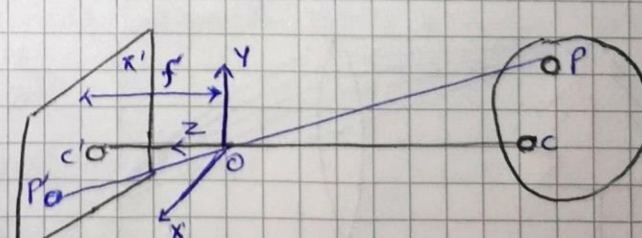
⑤ Perspective Distortion
(A) Problem for architectural photography: converging verticals.
(B) A sphere gets projected to:



(C) This distortion is not due to optics (lens). This problem was first pointed out by "Leonardo da Vinci".

(6) Modelling the Projection

The coordinate System



(A) We use the pinhole camera model as approximation of projection.
(B) We put the optical centre (focal point) at the origin (O)
(C) We put the image plane (π') in front of O.

Projection Equations

(A) Compute intersection of ray from $P = (x, y, z)^T$ to O and π'
(B) Derived using similar triangles.

(c) We get the projection by throwing out the last coordinate:

$$P = (x, y, z)^T \rightarrow \left(f' \frac{x}{z},\ f' \frac{y}{z}\right)^T = P'$$

} Is this a linear transformation? <u>No.</u> (division by z makes it non-linear)

★ Trick → A trick is to add 1 more coordinate.

$$\underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\text{Euclidian}} \rightarrow \underbrace{\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}}_{\text{Projective}}$$ → homogeneous image coordinates

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$ → homogeneous scene coordinates

converting from homogeneous coordinates:

(a) we have to make sure that our last coordinate is 1. i.e.

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \rightarrow \begin{pmatrix} \frac{x}{w} \\ \frac{y}{w} \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \rightarrow \begin{pmatrix} x/w \\ y/w \\ z/w \end{pmatrix}$$

⑦ Perspective Projection matrix

Projection is a matrix multiplication using homogeneous coordinates:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{f'} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z/f' \end{pmatrix} \Rightarrow \begin{pmatrix} f' \frac{x}{z} \\ f' \frac{y}{z} \end{pmatrix}$$

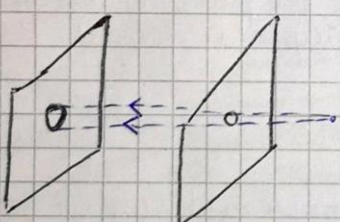
→ Dividing by the 3rd coordinate.

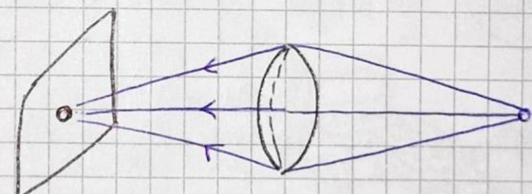
⑧ To get a clear image, we can shrink our aperture (as small as possible) but up to what extent? what happens when we do this?

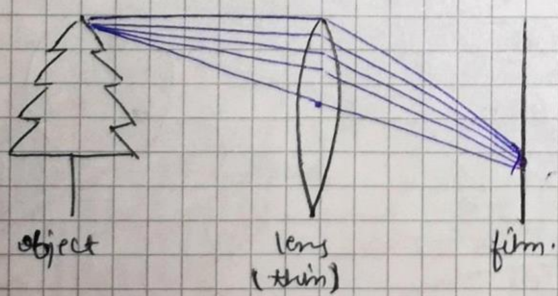
(a) less light gets through.

(b) ~~Distracting effects~~ Diffraction effects.

To overcome these effects, we use the concept of lenses.





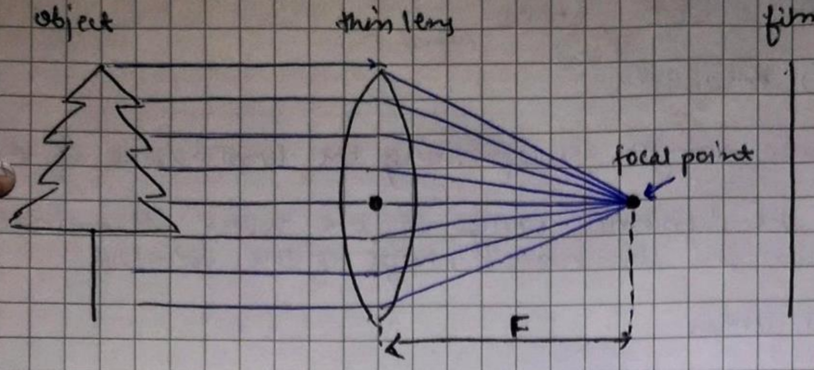


(A) A lens focusses the light onto the film.

(B) Rays passing through the center r not deviated.

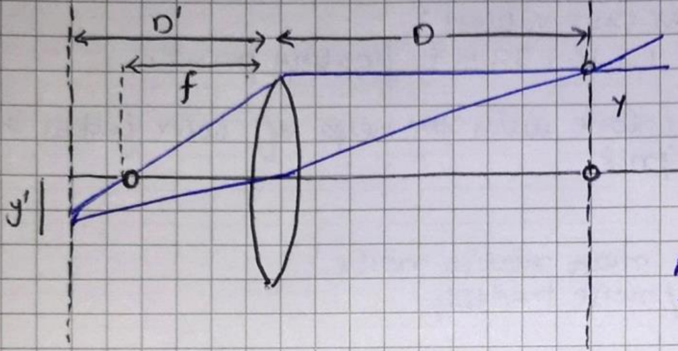
(c) There is a specific distance at which the objects are in focus.

(d) other points project to a circle of confusion in the image.



(A) All parallel rays converge to one point (focal point) on a plane located at a distance f (focal length) to the lens.

⑨ Thin lens formula.



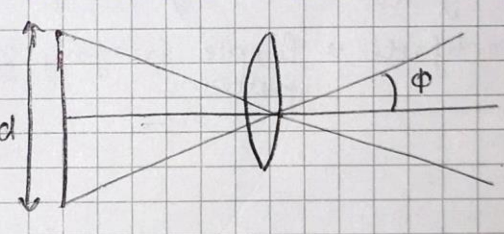
Similar triangles $\frac{y'}{y} = \frac{D'}{D}$

$$\frac{y'}{y} = \frac{D'-f}{f}$$

Any point satisfying the thin lens equation is in focus $\frac{1}{D} + \frac{1}{D'} = \frac{1}{f}$

⑩ We can control the depth of field by aperture size.
(A) If our aperture is more, objects close to us will be in focus. Just like in portrait photography.
(B) If the aperture is small, effect will not be very sharp. i.e. A smaller aperture increases the range in which the object is approximately in focus. Alternatively, a smaller aperture reduces the amount of light being received i.e. we need to increase the exposure.

⑪ Field of View (FoV)



(A) FoV depends upon the focal length and the size of the camera retina.

Smaller FoV = larger focal length $\phi = \tan^{-1}\left(\frac{d}{2f}\right)$

⑫ Chromatic abberation → Light is composed of several colours which usually we don't see. When a white light enters an optically active medium (eg. glass), all components gets refracted according to their wavelength. So, the refractive index depends upon wavelength of the light.

$$f = f(\lambda) : \quad \lambda\uparrow \sim f\downarrow$$

Vignetting → dark corners at the corners of an image

Radial distortion → (a) Caused by imperfect lenses
(b) Deviations are most noticeable for rays that pass the edge of the lens.

⑬ Digital cameras → (a) A digital camera replaces film with a sensor array.

(b) Each cell in the array is a light sensitive diode that converts photons to electrons.

(B) These are usually of 2 types : CCD , cmos

(14) High dynamic Range (HDR) images:

(a) CCD/cmos have a dynamic range for capturing the luminance of the scene → LDR cameras

(b) Problem → Sometimes, the dynamic range of the scene illumination is higher than the dynamic range of the camera.

(c) effects → Under exposed areas
Saturation areas.

(d) Creating HDR images → (a) Standard RGB images are encoded over 8 bit integers (255 values)
(b) HDR: 12, 16, 32 bit floating point.

(15) Color sensing in camera is basically done with the help of color filter array
eg. Bayer grid.

(16) Issues with digital cameras:

(A) Noise → Low light is where we most notice noise
→ Light sensitivity (ISO) / noise tradeoff

(B) Resolution → Requires higher quality lens
→ Noise issues

(c) In-camera Processing → Oversharpening can process halos.

(d) RAW vs. compressed → File size vs. quality tradeoff.

(e) Blooming effects → charge overflowing into neighbouring pixels.

(f) color artifacts → Purple fringing from microlenses, artifacts from Bayer's pattern.