lec 10 Stero Vision

Retina $\simeq 1000\,m^2$ Contain millions of photoreceptors
① rods → gray scale ② Cones → RGB
large proportion of our brain power is dedicated to process signals from eye

how do we see the world?
① Film in front of object → No Reasonable img
② Add barrier to Block most of rays
→ Reduce bluring → (حلها)
→ brarier Contain hole ⇒ يعني → aperture

Pinhole Camera model
① Capture Pencil of rays ⇒ all rays path through Single Point
② Point Called Center of Projection او optical Center
③ Image formed on Image plane

لو خلينا الaperture صغير اوي ⇒ ① less light get through (increase exposure)
② Diffraction effect

↓ الحل

add lens in aperature
① ⇒ lens focus light on the film
② " Rays passing through Center are not deviated

③ All rays Parallel to optical axis Converge at focal Point

/ /

Thin lens equation

$$\frac{A}{B} = \frac{e}{z} \Rightarrow \frac{1}{F} = \frac{1}{z} + \frac{1}{e}$$

Focal length

distance between Center and object

distance Center and image / film

any point Satisfy this equation
⇒ is in Focus
others blur circle

This formula used to estimate the distance to object (Depth from focus)

dependence of the apparent size of objects on their distance from observer is known as Prespective
gives us ← strong depth cues → we can perceive 3D Scene by view its 2D representation
when view 3D Scenes
↳ it may be mislead by perception

* Camera doesn't measure distances but angle (bearing sensor)

Image plane is usually represented in front So that img preserves the same orientation (Not flipped)

For Convenience

ALADIB

/ /

Perspective equations

$$\frac{x}{f} = \frac{X_c}{Z_c} \Rightarrow \boxed{x = \frac{f X_c}{Z_c}} \qquad y/f = \frac{y_c}{Z_c} = \boxed{\frac{f y_c}{Z_c}}$$

from Camera frame to pixel Coordination (CpI)

local img plane (x, y) pixel Coordinate (u, v)

① pixel Coordinate of Camera optical Center (u_0, v_0)

② Scale factor of pixels size in k_u, k_v

$$\boxed{u = u_0 + \frac{k_u f X_c}{Z_c}} \qquad \boxed{v = v_0 + \frac{k_v f Y_c}{Z_c}}$$

homogenous Coordinates $\Rightarrow$ linear mapping from 2d $\rightarrow$ 3d

$\boxed{\lambda = 1}$

Calculate Intrinsic Matrix

$$\boxed{k \begin{bmatrix} X_c \\ y_c \\ Z_c \end{bmatrix}}$$ Projection matrix

---

From Camera Frame to world frame

extrinsic Matrix (cwe)

$$k[R|T]\begin{bmatrix} X_w \\ y_w \\ Z_w \\ 1 \end{bmatrix} \Bigg\} \text{projection Matrix}$$

/ /

Radial distortion

$$\begin{bmatrix} u_d \\ v_d \end{bmatrix} = (1 + k_1 r^2) \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}$$

k_1 → radial distortion Parameter

$$r^2 = (u - u_0)^2 + (v - v_0)^2$$

Camera Collibiration (Stero Camera Collibiration)

⇒ measure all unknown parameters to form Camera model (intrinsic, extirsnic)

$\boxed{\text{pixel Coordinates}}$ → Intirnisic Parameters بتعرف من

Impossible to Capture 3d structure from a single view
⇓
لذا
1. observe scene from 2 differnt points
2. Solve the intersection of rays and relover 3D

how do we measure distance with Camera?
1. from (Stero vision)
   2 Cameras with known relative position, orientation
2. Structure from motion
   Single moving Camera, both 3D structure and Camera motion Can be estimated up to a Scale

/ /

Stero Vision - "Simplest Case"

Ideal Case ⇒ Both Cameras are Identical, aligned at horizontal axis

Base line ⇒ distance between optical Center of 2 Camera

Disparity ⇒ difference in image location of projection of 3D Point in 2 image planes

$$\underbrace{Z_P}_{\text{depth}} = \frac{\overbrace{bf}^{\text{Baseline}}}{\underbrace{u_l - u_r}_{\text{disparity}}}$$

- Depth is inversely proportional to disparity

$Z_P \propto \frac{1}{u_l - u_r}$ ⇒ Foreground have bigger disparity than Background object

Foreground more Brighter

- Disparity is $\propto$ Stero Base line (b)

Small Basline ⇒ more unCertain our estimate Path

b increase ⟷ object may apper in one Camera

Projections of a 3D Point onto left, right Stero images are Called Correspondance Pair

light Color → large disparity

الأقرب Foreground lighter than Background

Disparity map → hold disparity value of each pixel } lighter ↓ large disparity ↓ object close to camera

Stereo Vision (General Case)

محتاج

using Collibration methods {
① relative pose between 2 cameras — rotation, translation
② focal length, image center, radial distoration

Epipolar Geometry $\xrightarrow[\text{مشكلة}]{\text{حل}}$ Corresspondence problem

make Image search (1D) with tolerance

Epipolar plane = defined by 3D point (P) + optical center

Epipolar rectification

Determine transformation of each img plane so that pairs conjugate epipolar lines become colliner, (||) to one of img axis

---

3D ReConstruction ⟹ tringulation

triangulate correspondent to get

Stero Vision Summary

/ /

1. Stero Camera Callibration → Compute Camera relative pose
2. Epipolar rechification → align image, epipolar lines
3. Search Corresspondances

output: triangulation, disparity map