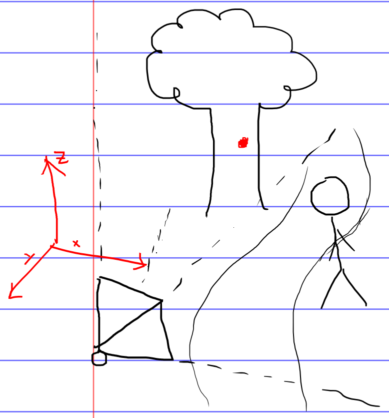
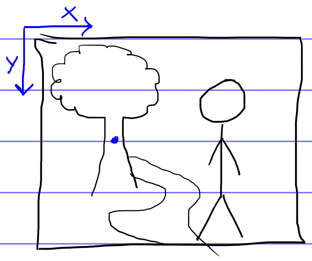
Lecture 3: Image Geometry

Topics:
- Frames of reference
- Image planes
- Projection + transformations
- Homogeneous coordinates
- Two-camera geometry
- Epipolar geometry
- Depth from disparity

3D Scene ⟶ 2D Image



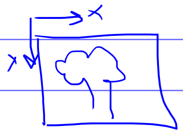


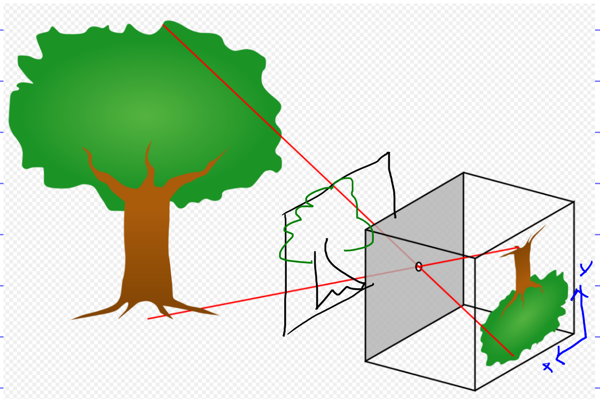
How does [x,y,z] relate to [x,y]?

→ How do cameras work?

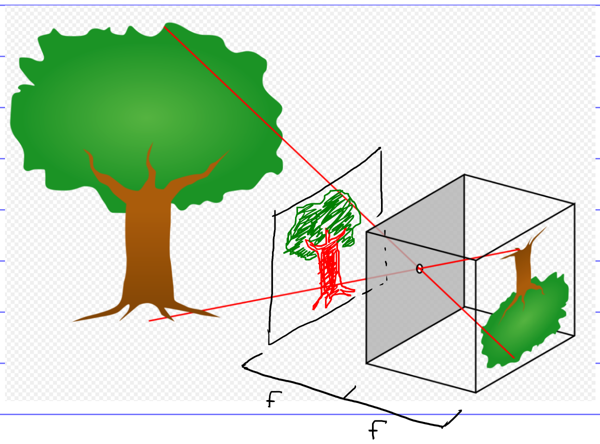
Pinhole camera model:

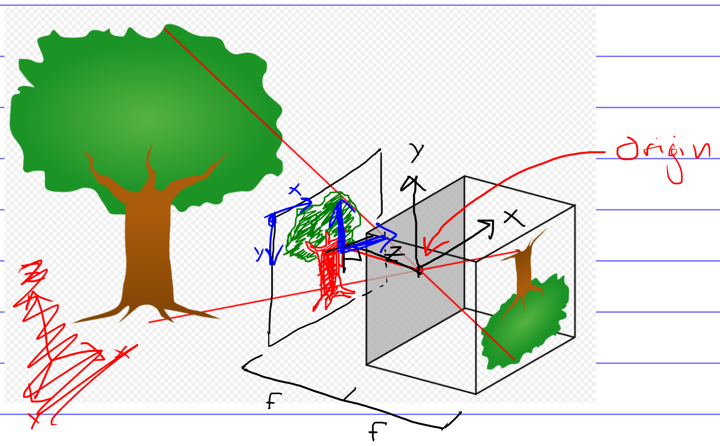
+ Simple and accurate
- Does not account for focus or depth of field





- Produces image flipped 180° — two ways to mitigate this
  - Redefine x & y
  - Redefine image plane

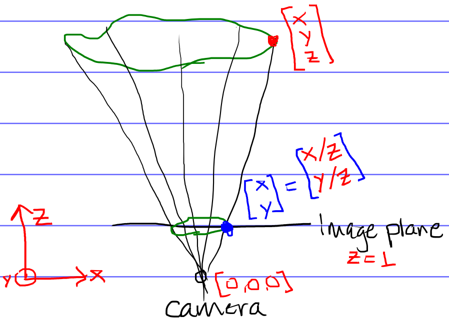




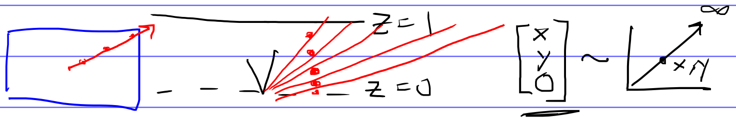
3D ⟶ 2D

1. Choose [0,0,0] → pinhole
2. Choose (z, y, x axes) → z: out of pinhole towards image plane; x,y = image x,y
3. Choose f = 1
4. Choose (y, x axes) → Same as x, y

Then image formation is linear projection onto z = 1

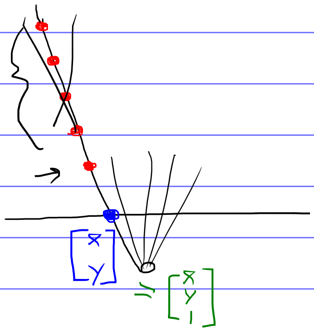


$$\begin{bmatrix}x\\y\\z\end{bmatrix} \xrightarrow{\text{proj}} \begin{bmatrix}x/z\\y/z\\1\end{bmatrix} \xrightarrow{\text{2D}} \begin{bmatrix}x/z\\y/z\end{bmatrix}$$



Notes:

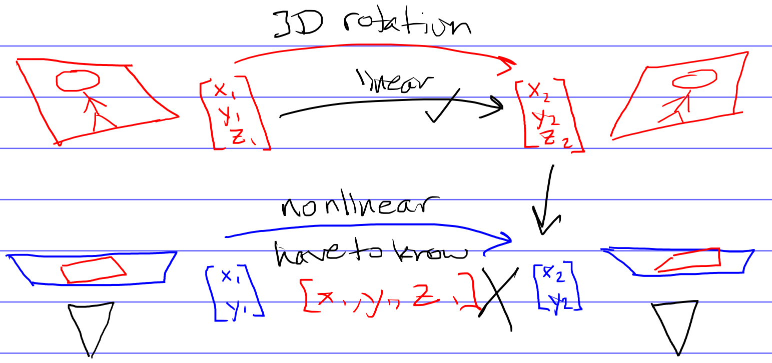
- If z=0? → Ideal points (infinity in a direction)
- If z<0? → N/A
- Unique? → Occlusion

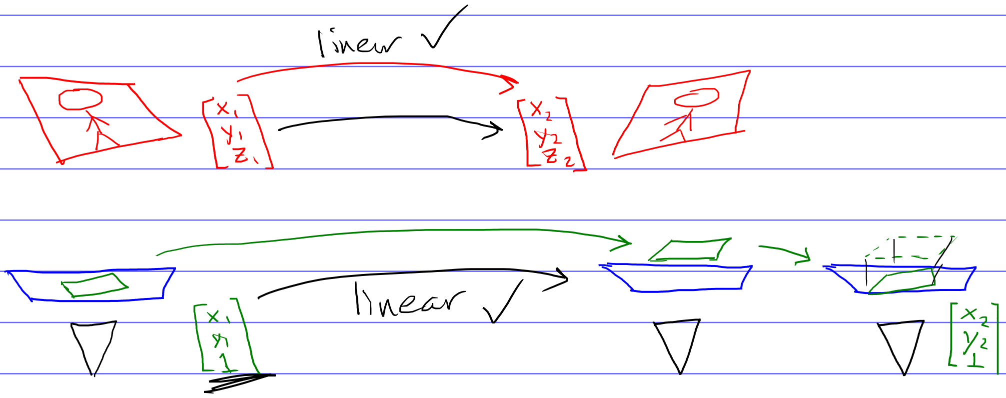


Homogeneous Coordinates

- We almost never know $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, only $\begin{bmatrix} x \\ y \end{bmatrix}$
- We want to be able to represent 3D transformations in our 2D image space
- Solution: homogeneous coordinates.

$$A * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x'/z' \\ y'/z' \\ 1 \end{bmatrix}$$





Hom. coordinates are useful for performing 3D transformations in the image plane.

Group	Examples	DoF	Matrix	Invariants
Projective	Homography, perspective warp	8	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$	Collinearity, intersection, tangency, cross ratio (ratio of ratio of lengths)
Affine	Skew, stretch	6	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	All of the above, plus parallelism, ratio of areas, ratio of lengths among collinear points (e.g. midpoints), linear combinations of vectors (e.g. centroids)
Similarity	Uniform scaling	4	$\begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$	All of the above, plus ratio of lengths, angles
Euclidean	Translation, rotation	3	$\begin{bmatrix} \cos\theta & -\sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$	All of the above, plus length, area

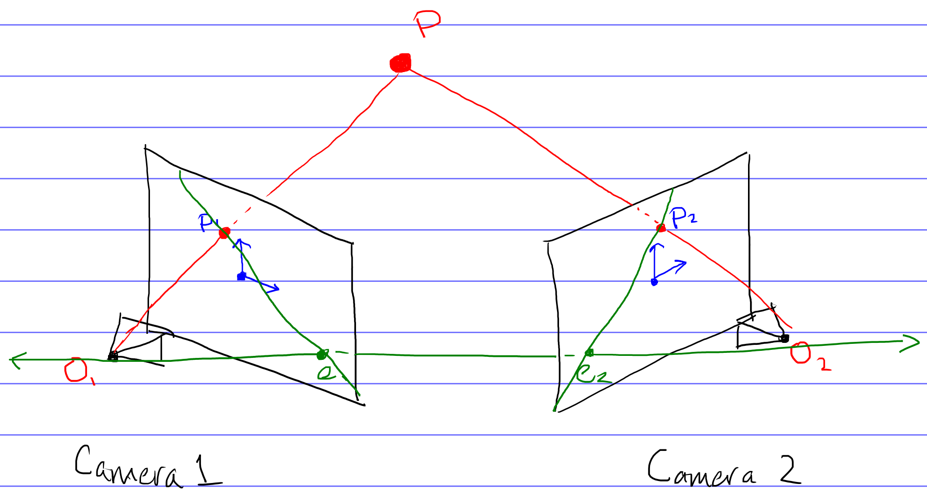
(You do not need to memorize these)

Two-camera (Epipolar) Geometry

- Also called stereo geometry
- Yields information about scene depth (z)
- Also useful when we have one moving camera

- Note: We now have <u>three</u> frames of ref. and must be careful with our coordinates

Basic setup:



Terms:
- Baseline - Line $\overline{O_1 O_2}$

- Epipoles - Intersection of baseline and image plane

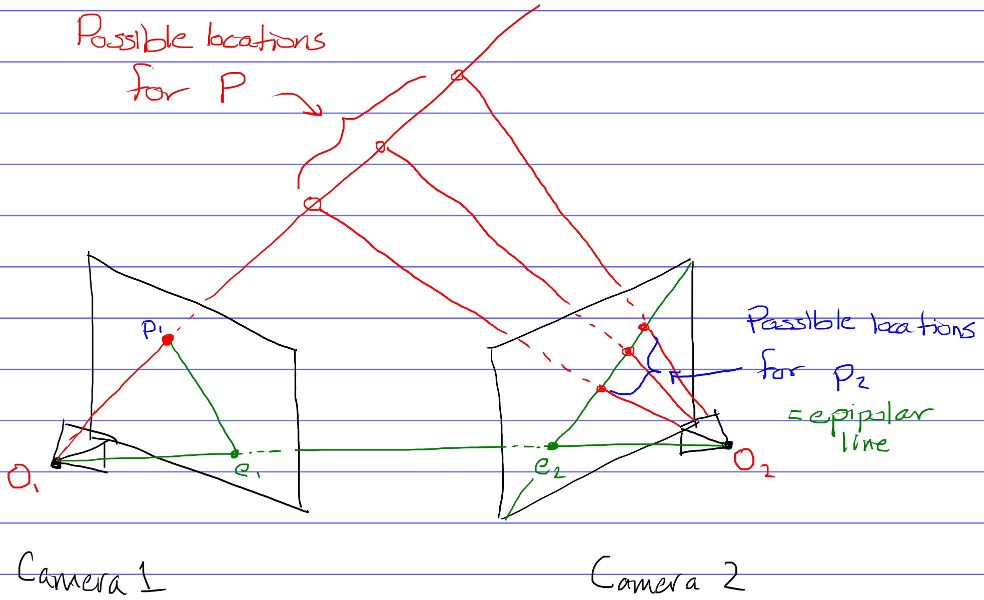
- Epipolar plane (containing P) - plane $O_1 O_2 P$

- Epipolar lines - Intersection of ep. plane & image plane

Q: Suppose we know O_1, O_2, and p_1. We don't know P. What do we know about p_2?

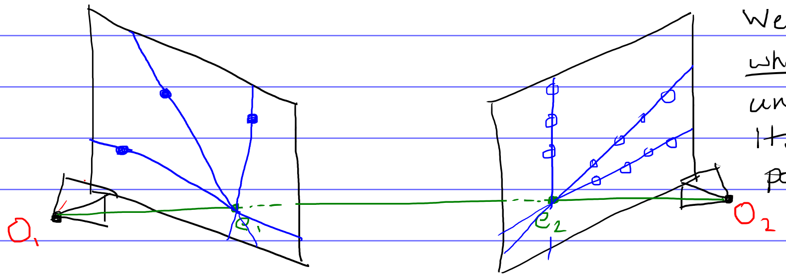
→ Useful in object tracking, scene reconstruction

A: We know p_2 lies on the epipolar line



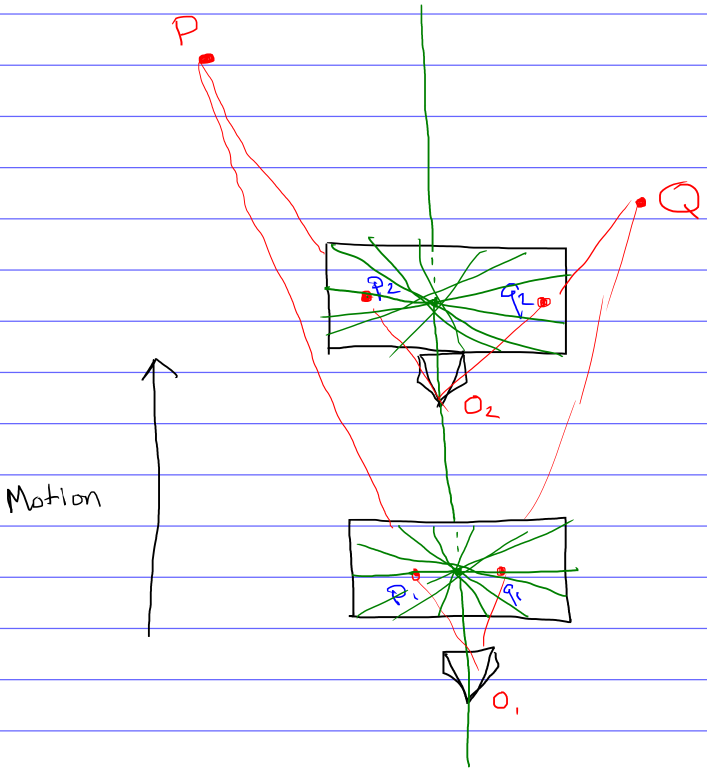
Intuition: Ep. line $\overrightarrow{e_1 p_1}$ is $\overrightarrow{O_2 P}$ projected onto image plane 1, and $\overrightarrow{e_2 p_2}$ is just $\overrightarrow{O_1 P}$ projected onto image plane 2.

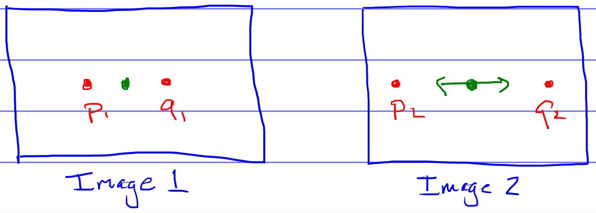
Takeaway: When we move from one camera position O_1 to another O_2, every point in our image must appear along an epipolar line.



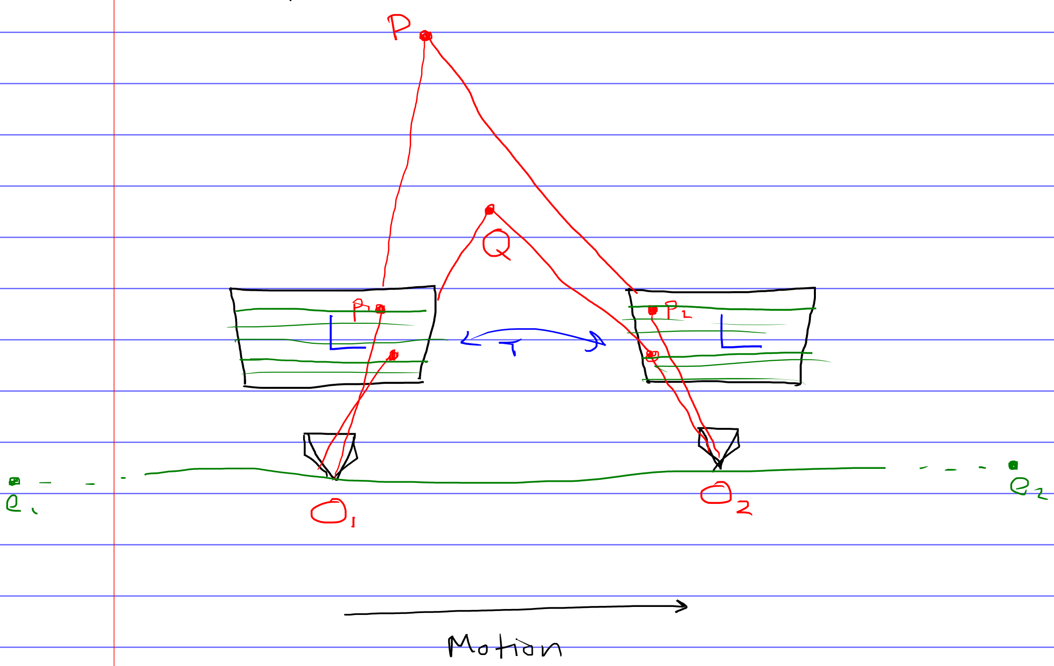
We don't know where on the line unless we know its real-world position.

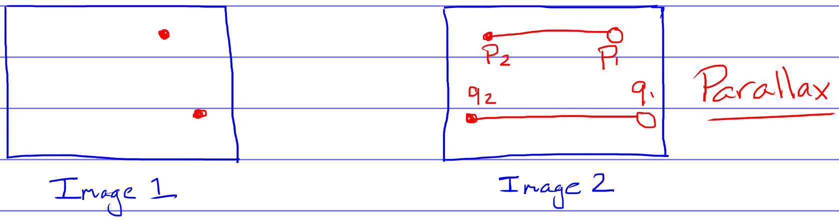
Example: Forward motion





Example: Lateral Movement



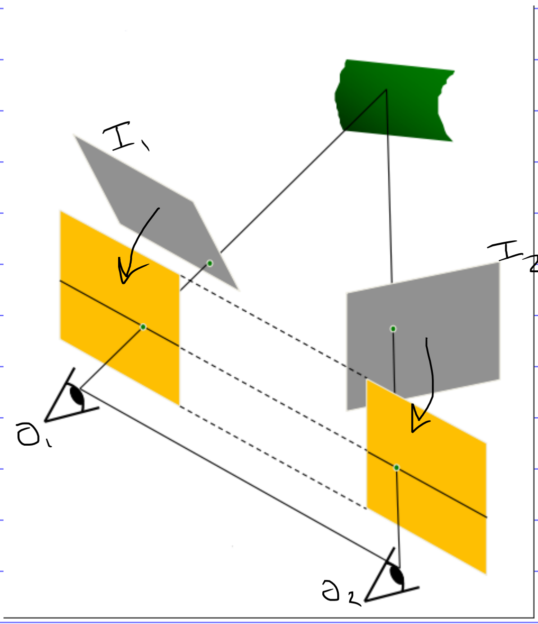


Q: How do we work in stereo algebraically?

A: With difficulty. ^^

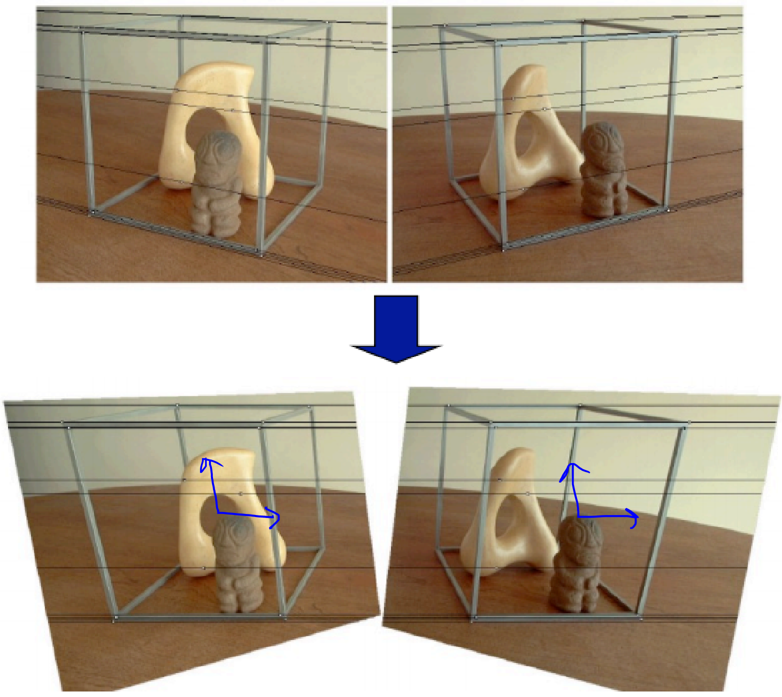
Real answer: Whenever possible, we work with purely lateral motion.

→ Rectification: Standardize our image plane w/ homographies



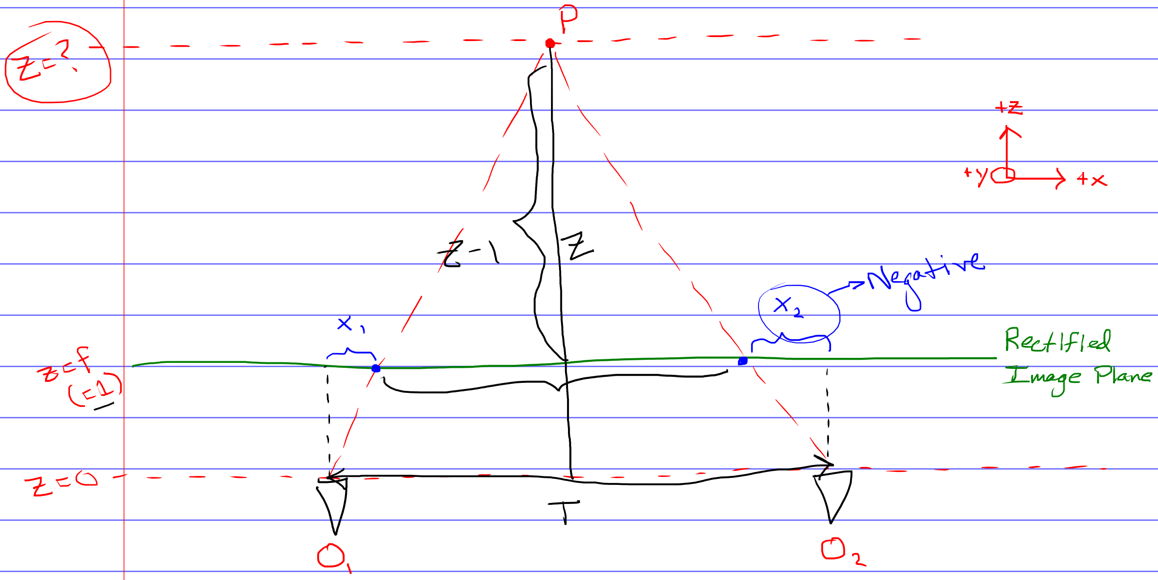
1. Find a new, common image plane
2. Apply a homography on each image to project the scene onto the new plane
3. Use our horizontal epipolar lines to learn about the scene.

Note: We do not move O_1 & O_2. But we do change the direction they are facing



Application: Depth from Disparity

Consider this scene: (viewed from above)



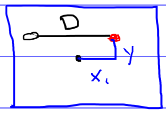


Image 1

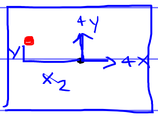


Image 2

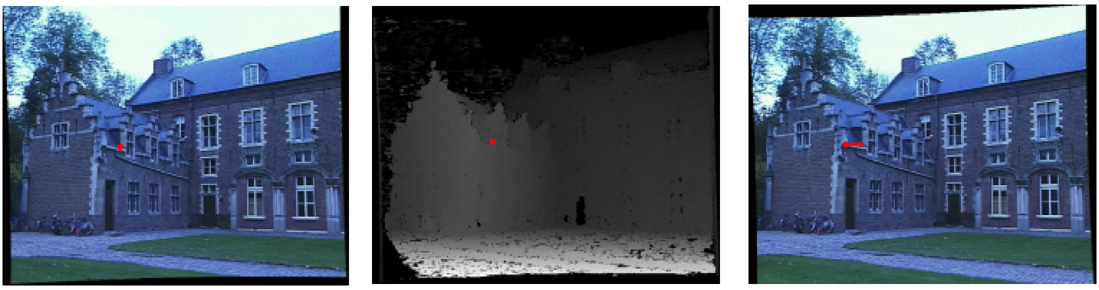
$$\frac{T}{Z} = \frac{T - x_1 + x_2}{Z - f} \longrightarrow \cancel{T}(\cancel{Z}-f) = \cancel{Z}(\cancel{T} - x_1 + x_2)$$

$$+fT = Z(+x_1 - x_2)$$

$$Z = \frac{fT}{\underbrace{x_1 - x_2}_{\text{Disparity}}}$$

Depth

Image 1 Disparity Image 2



≈ Depth