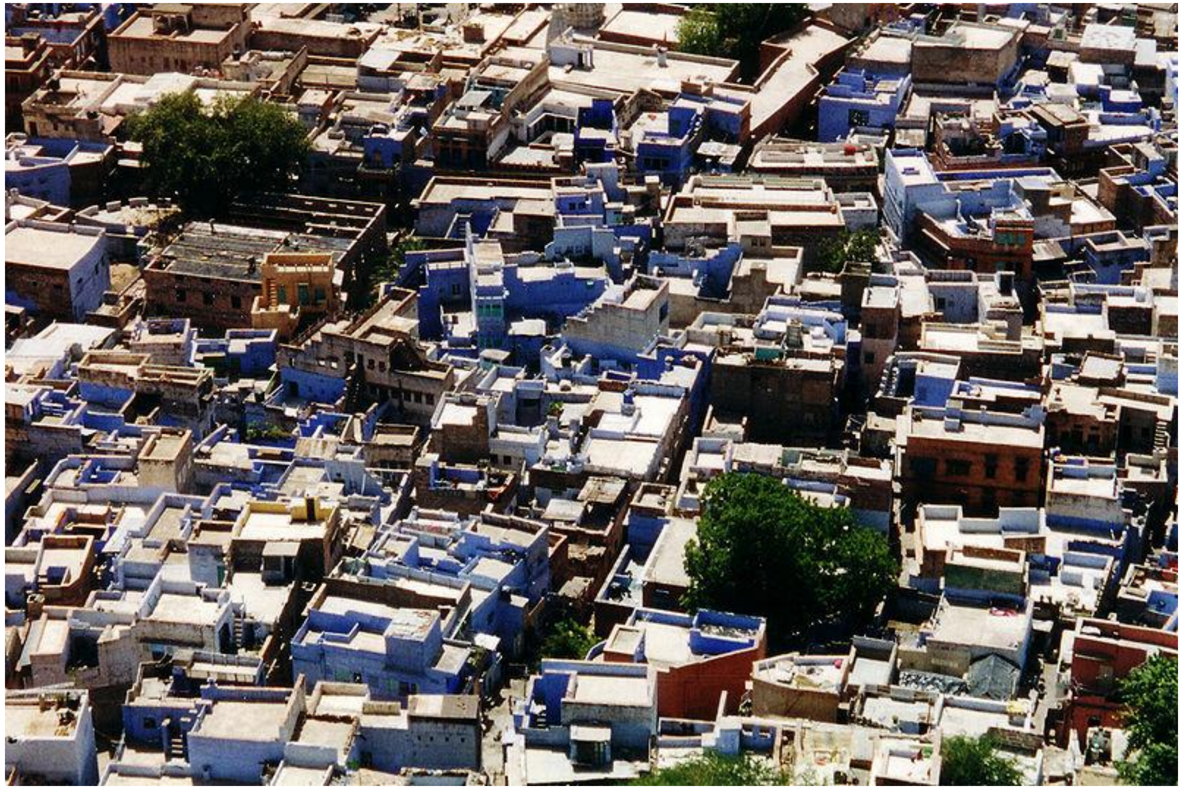
Miniature faking

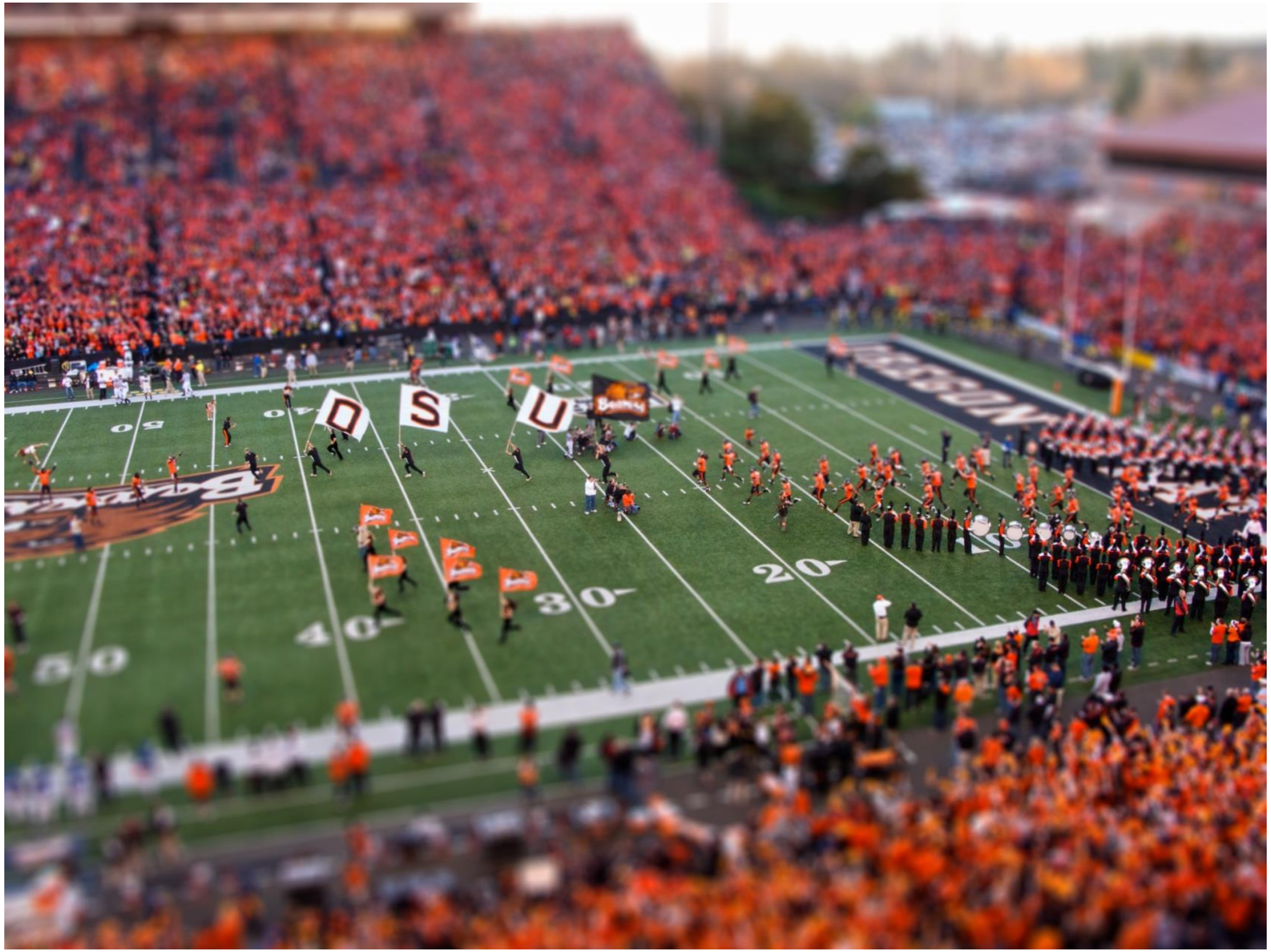
In close-up photo, the depth of field is limited.

http://en.wikipedia.org/wiki/File:Jodhpur_tilt_shift.jpg

Miniature faking



Miniature faking

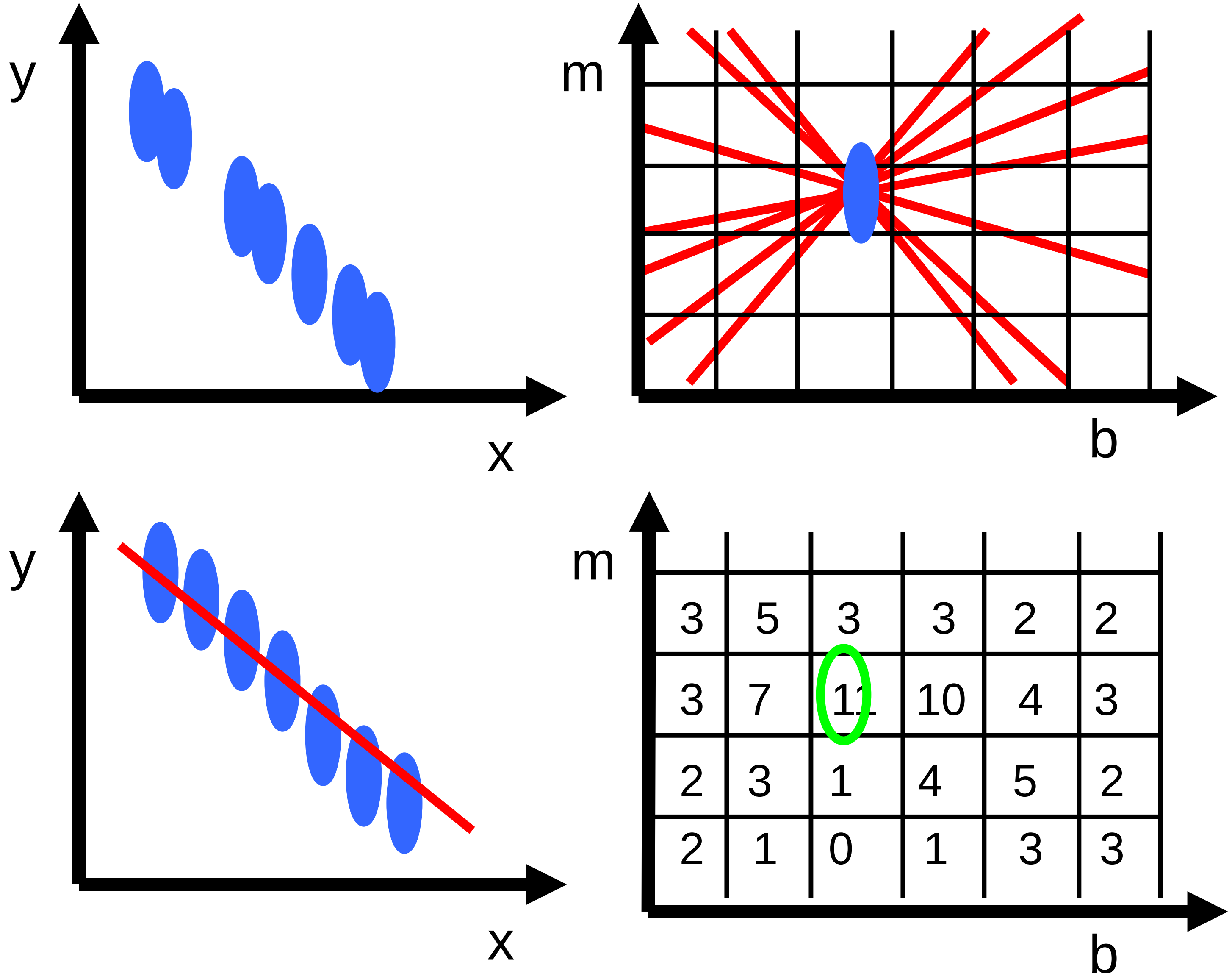


http://en.wikipedia.org/wiki/File:Oregon_State_Beavers_Tilt-Shift_Miniature_Greg_Keene.jpg

Review

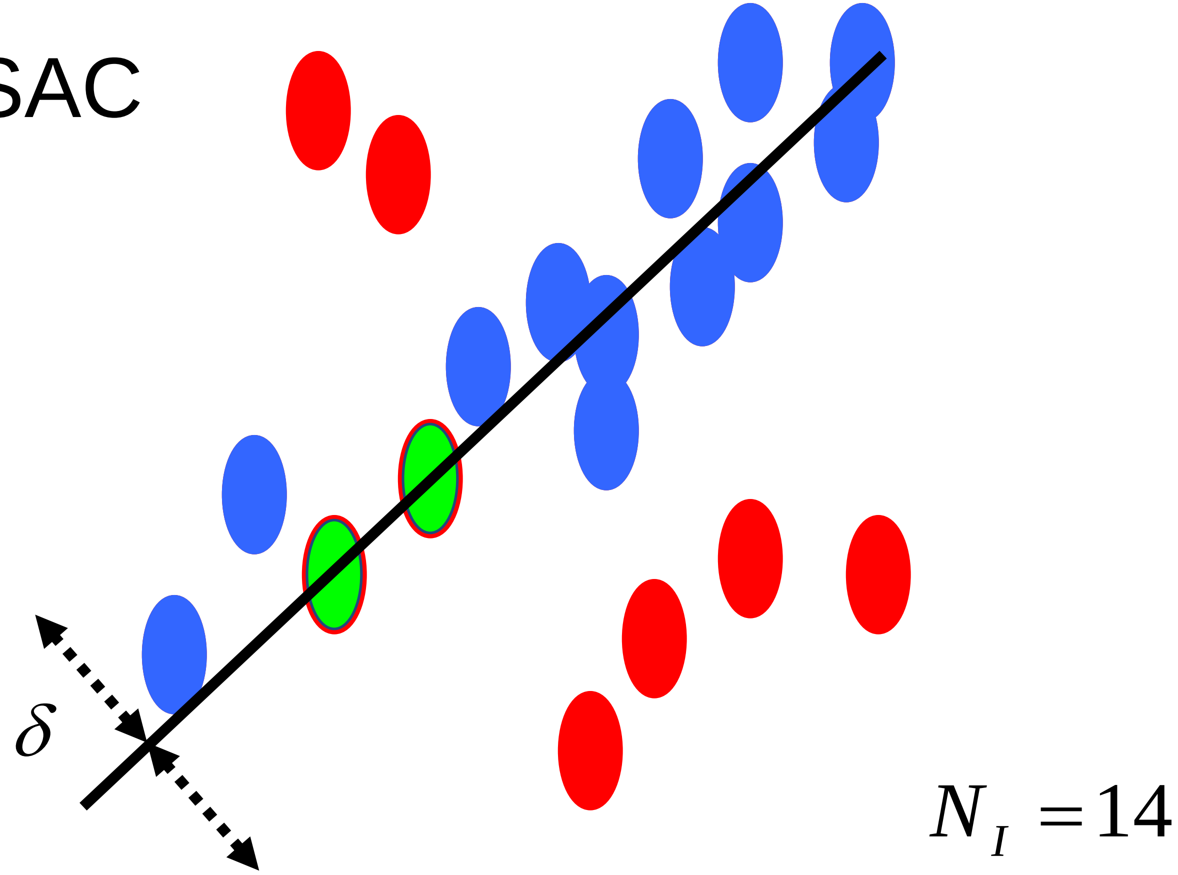
- Previous section:
  - Model fitting and outlier rejection

Review: Hough transform



Slide from S. Savarese

Review: RANSAC

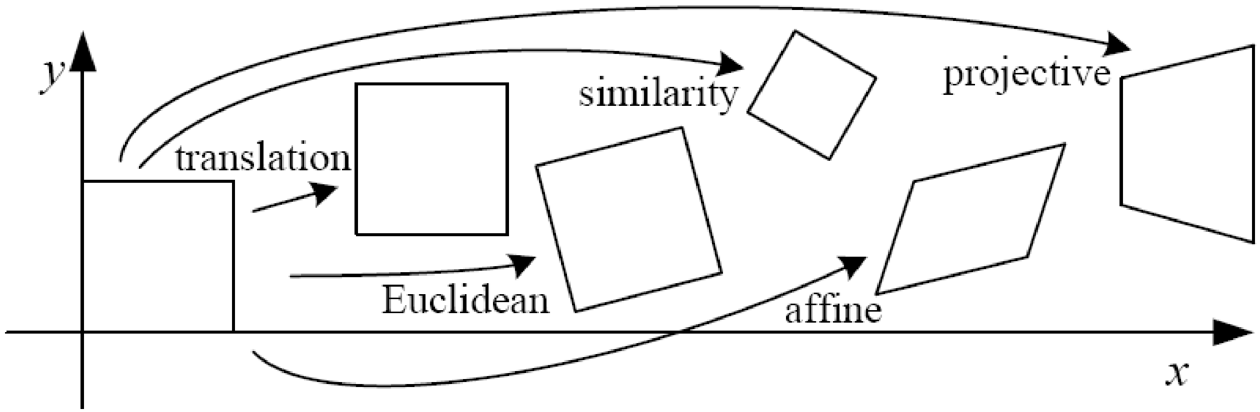


Algorithm:

1. **Sample** (randomly) the number of points required to fit the model (#=2)
2. **Solve** for model parameters using samples
3. **Score** by the fraction of inliers within a preset threshold of the model

Repeat 1-3 until the best model is found with high confidence

Review: 2D image transformations



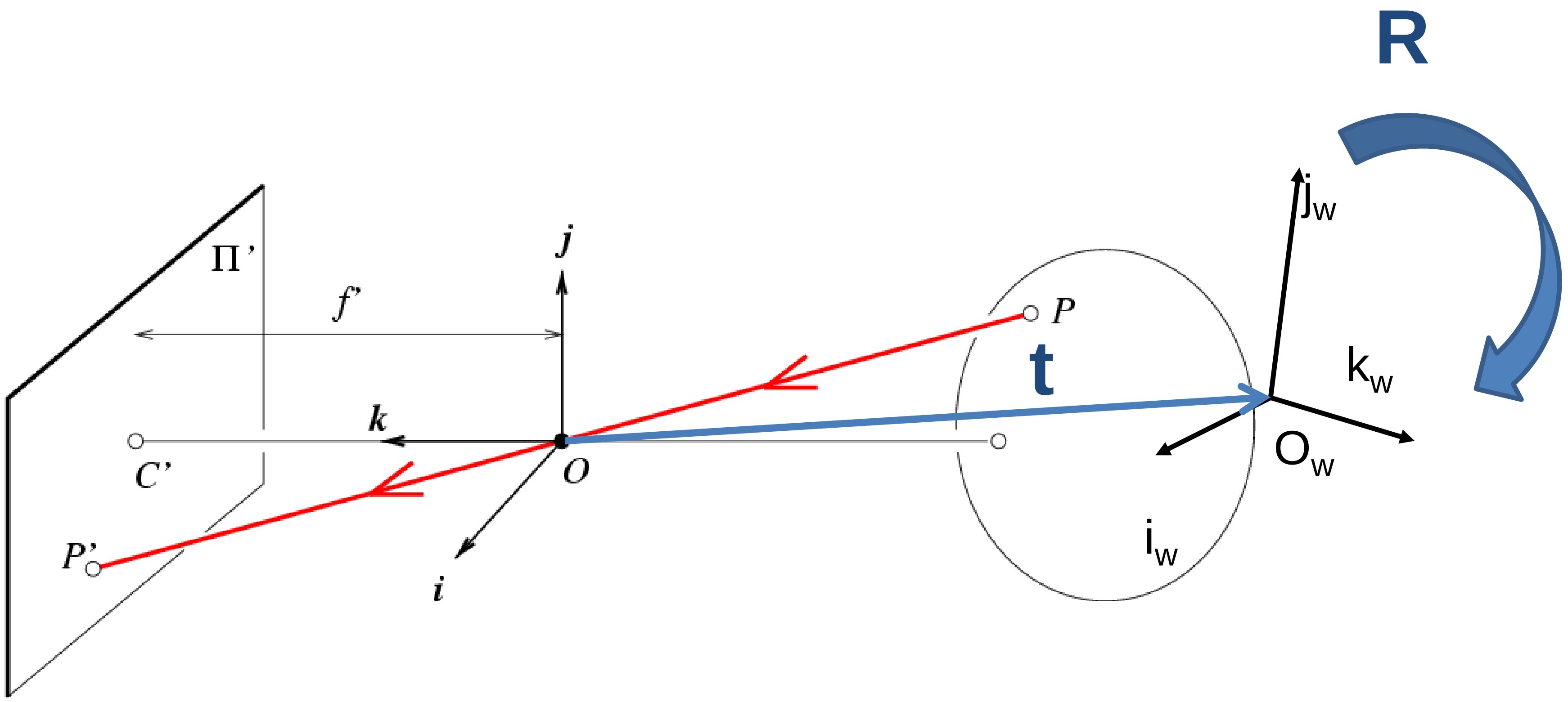
Name	Matrix	# D.O.F.	Preserves:	Icon
translation	$\left[\, \boldsymbol{I} \mid \boldsymbol{t} \,\right]_{2\times 3}$	2	orientation $+ \cdots$	
rigid (Euclidean)	$\left[\, \boldsymbol{R} \mid \boldsymbol{t} \,\right]_{2\times 3}$	3	lengths $+ \cdots$	
similarity	$\left[\, s\boldsymbol{R} \mid \boldsymbol{t} \,\right]_{2\times 3}$	4	angles $+ \cdots$	
affine	$\left[\, \boldsymbol{A} \,\right]_{2\times 3}$	6	parallelism $+ \cdots$	
projective	$\left[\, \tilde{\boldsymbol{H}} \,\right]_{3\times 3}$	8	straight lines	

Szeliski 2.1

This section – multiple views

- Today – Intro to multiple views and Stereo. Camera Calibration (if we have time).
- Next Lecture – Epipolar Geometry and Fundamental Matrix. Stereo Matching (if there is time).
- Both lectures are the core of what you need for project 3.

Recap: Oriented and Translated Camera



Recap: Degrees of freedom

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

5 6

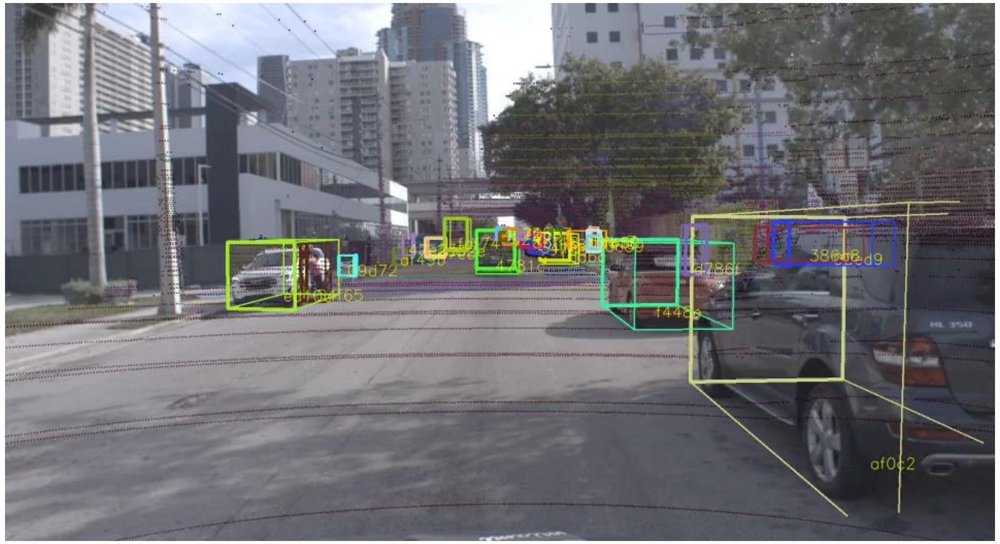
$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

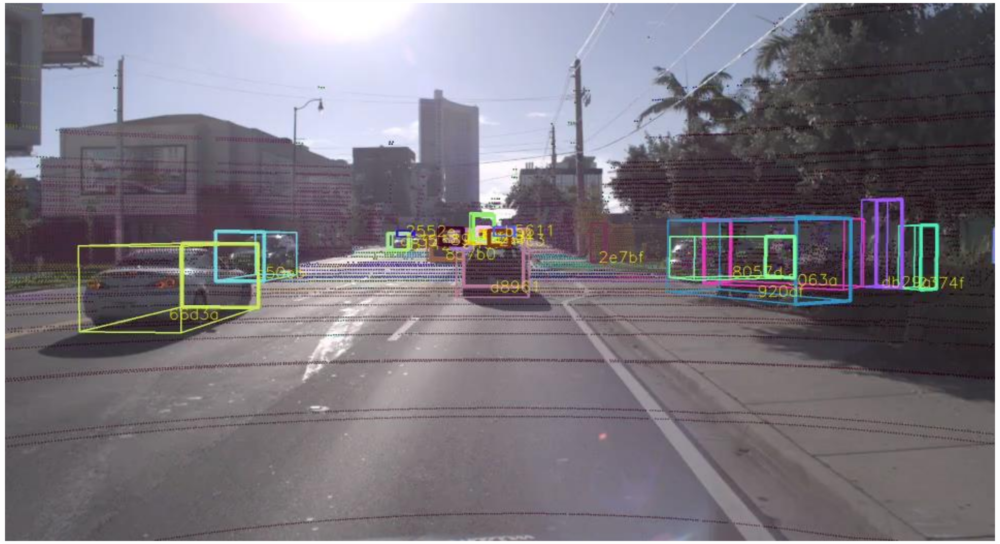
This Lecture: How to calibrate the camera?

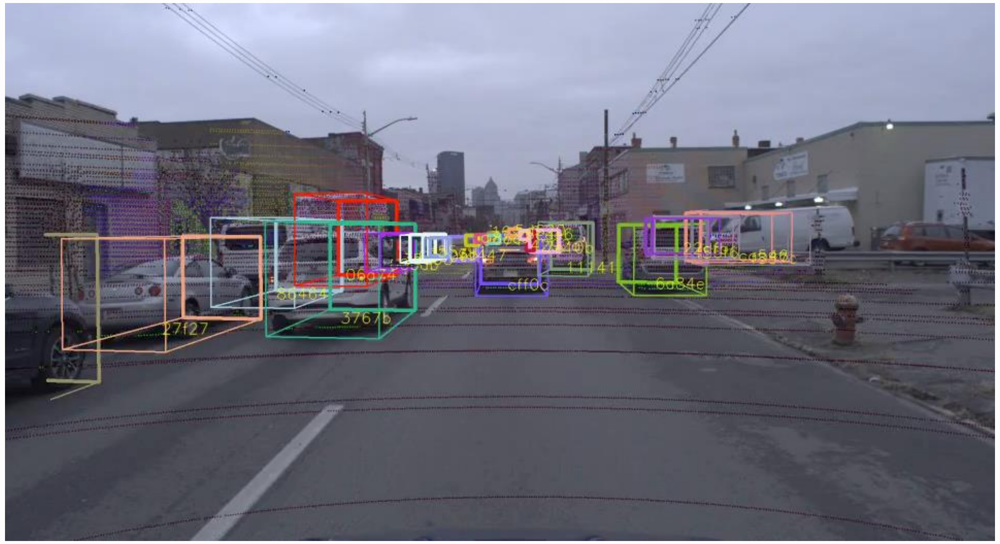
$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

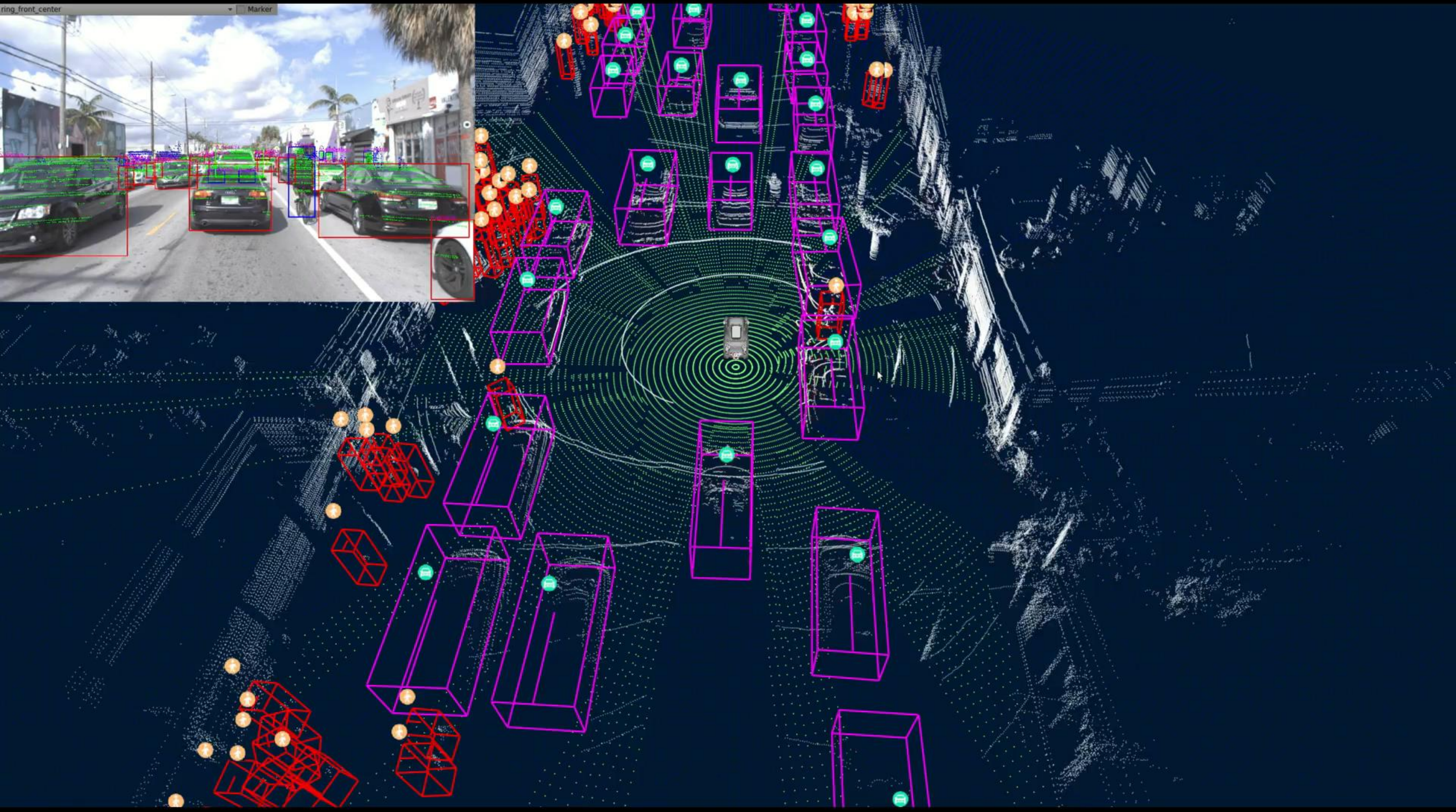
$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

What can we do with camera calibration?





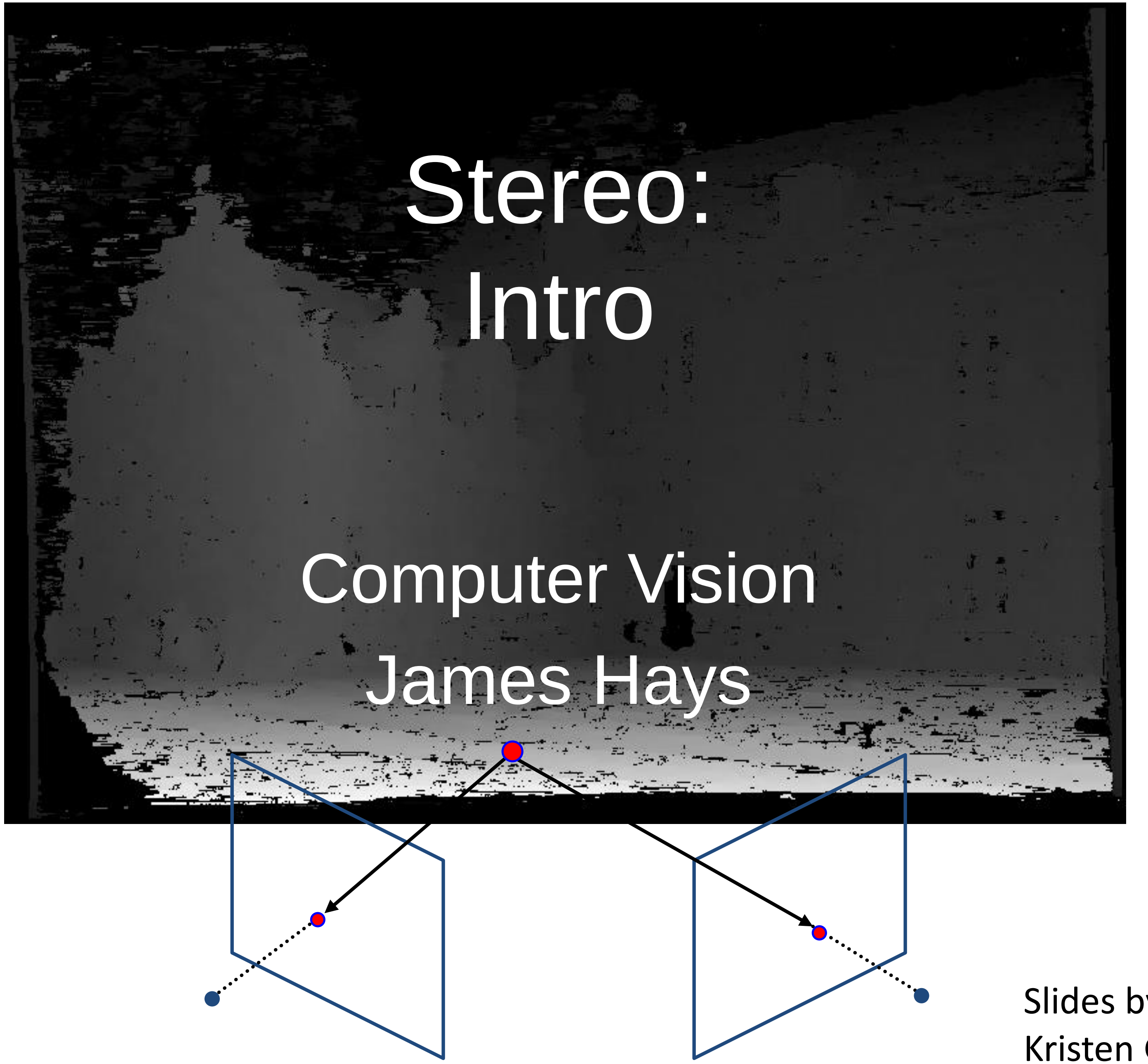



ring_front_center
Marker

Stereo: Intro

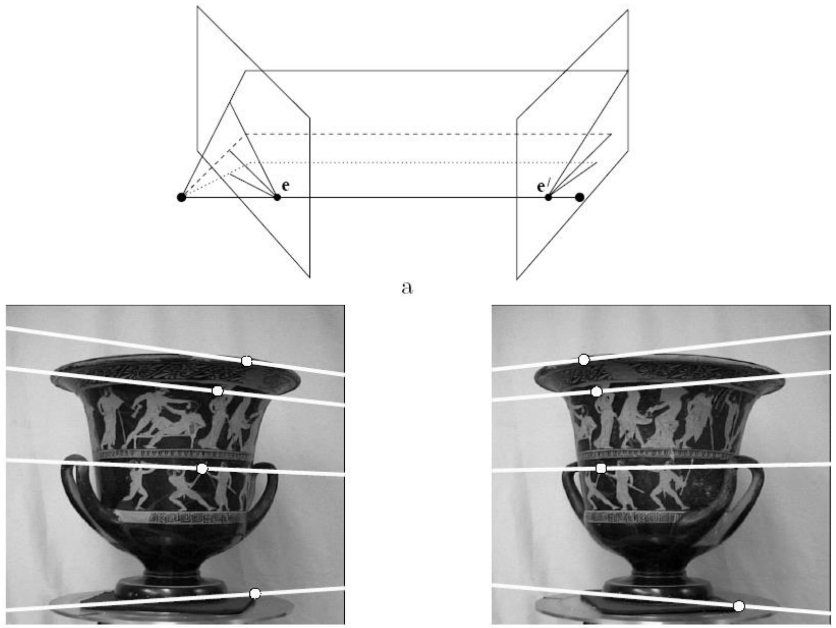
Computer Vision

James Hays



Slides by
Kristen Grauman

Multiple views



Hartley and Zisserman

stereo vision
structure from motion
optical flow

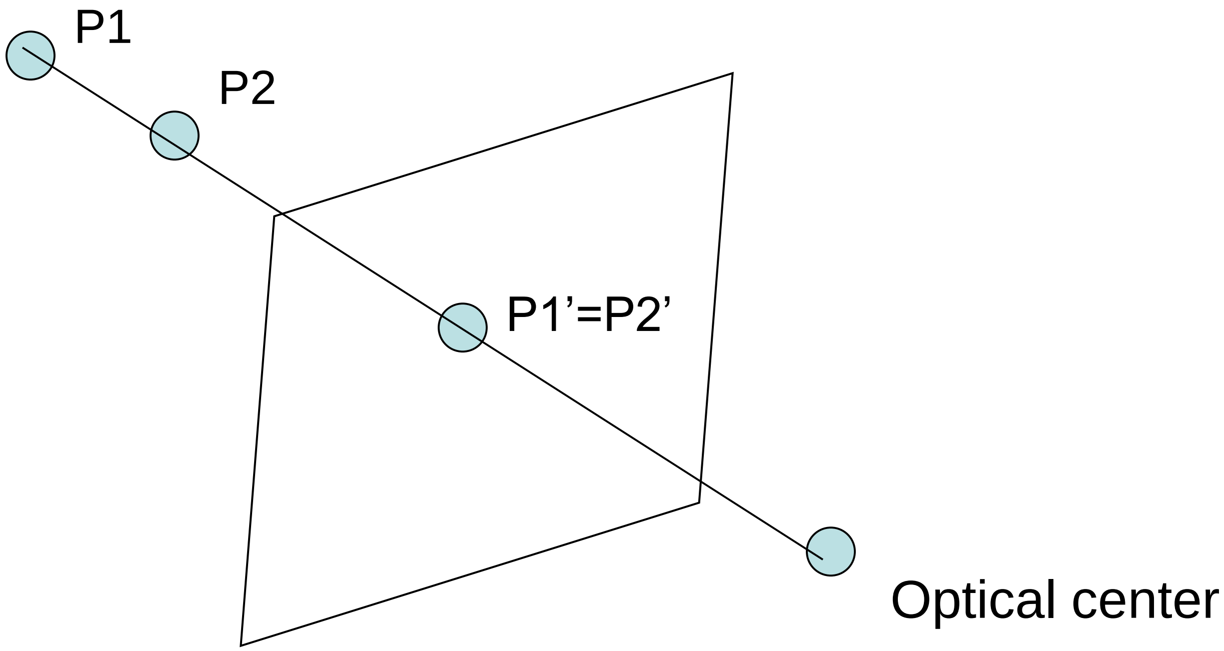
Why multiple views?

- Structure and depth are inherently ambiguous from single views.

Images from Lana Lazebnik

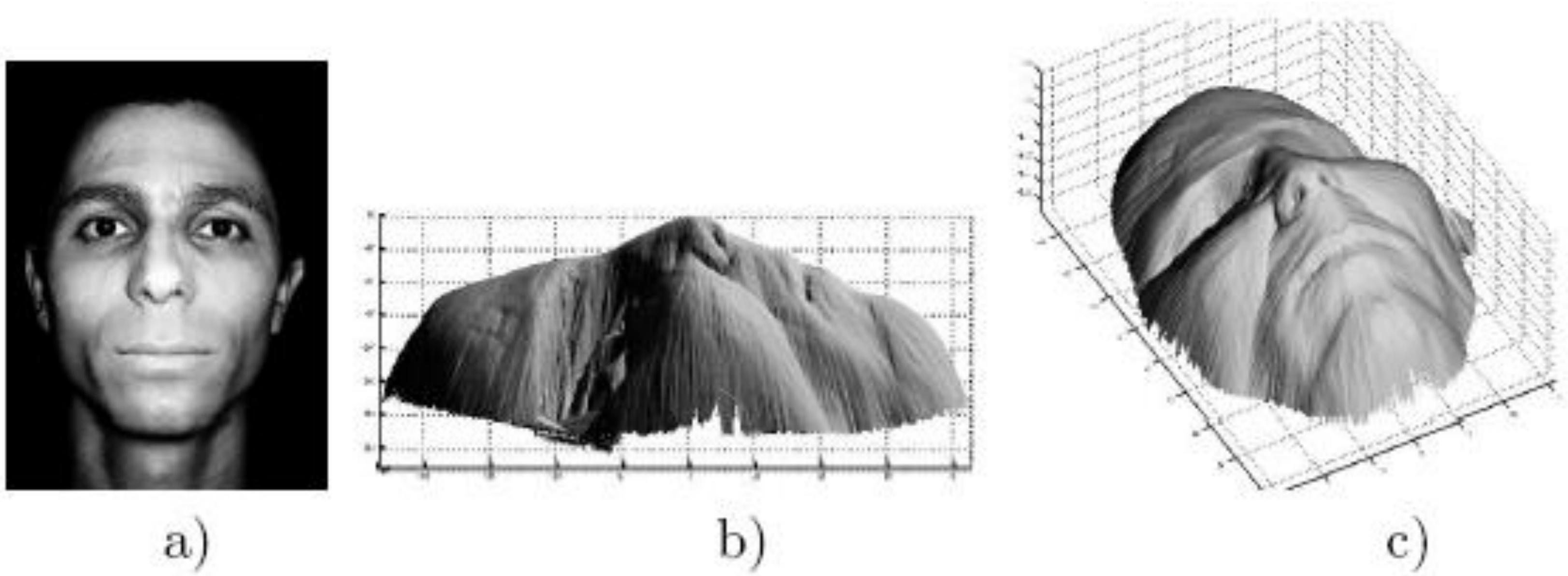
Why multiple views?

- Structure and depth are inherently ambiguous from single views.



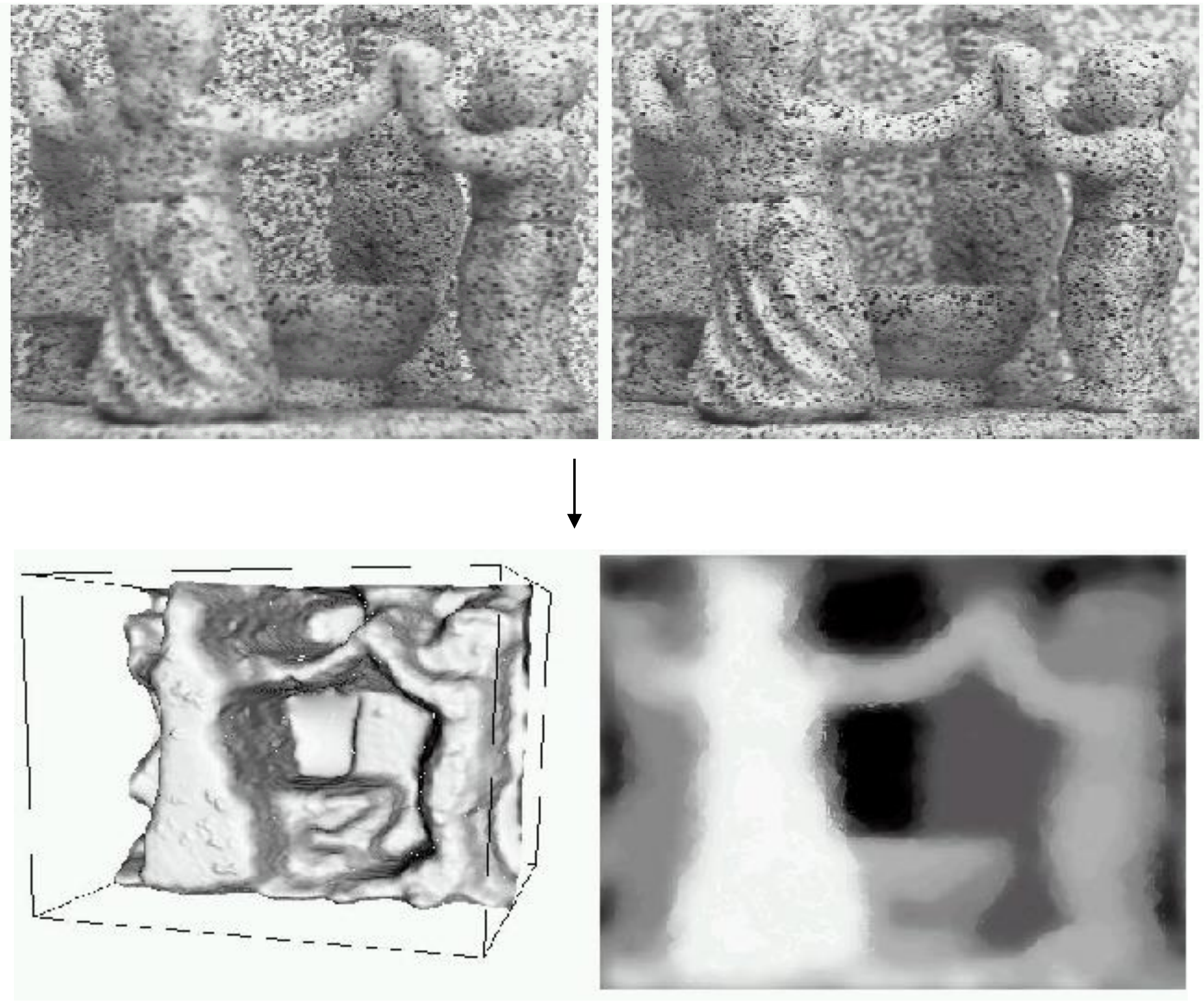
- What cues help us to perceive 3d shape and depth?

Shading



[Figure from Prados & Faugeras 2006]

Focus/defocus

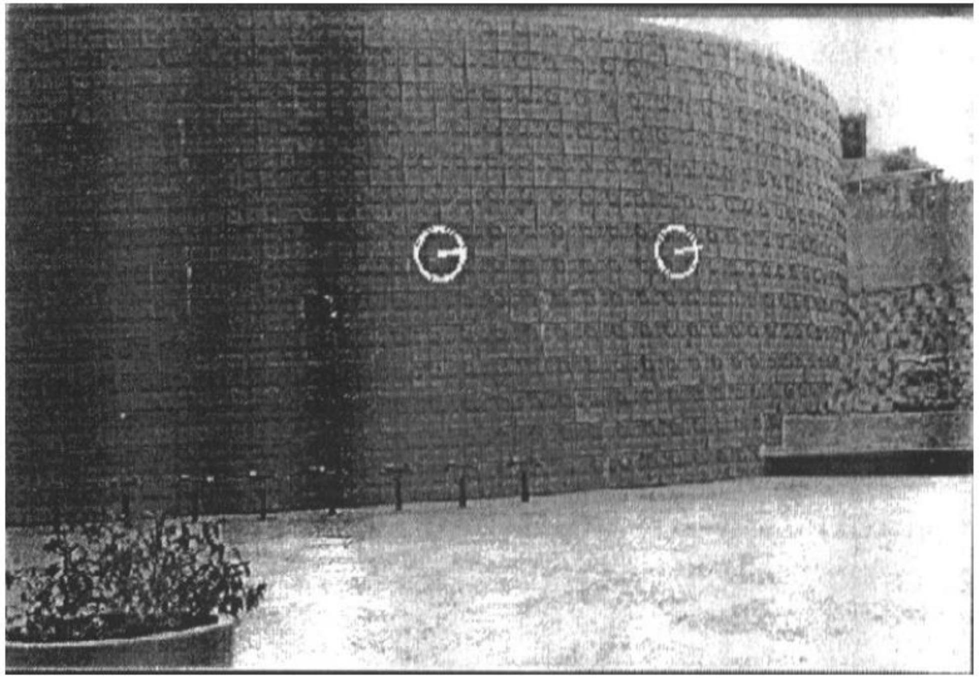


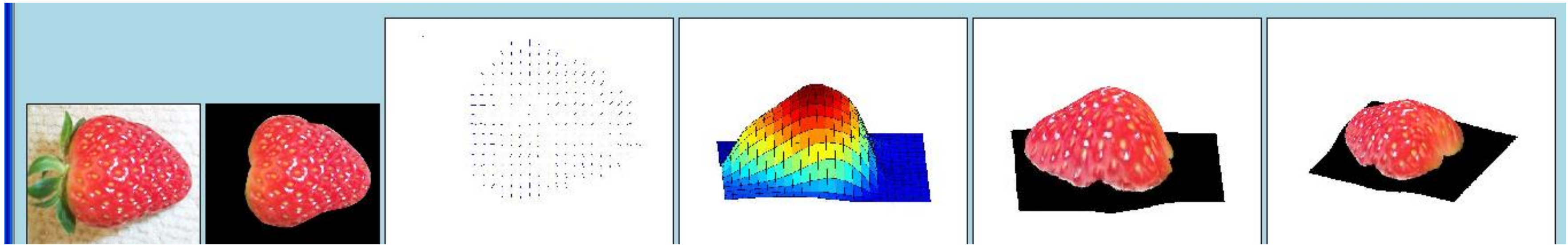
Images from same point of view, different camera parameters

3d shape / depth estimates

[figs from H. Jin and P. Favaro, 2002]

Texture



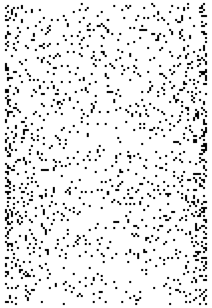


[From A.M. Loh. The recovery of 3-D structure using visual texture patterns. PhD thesis]

Perspective effects

Image credit: S. Seitz

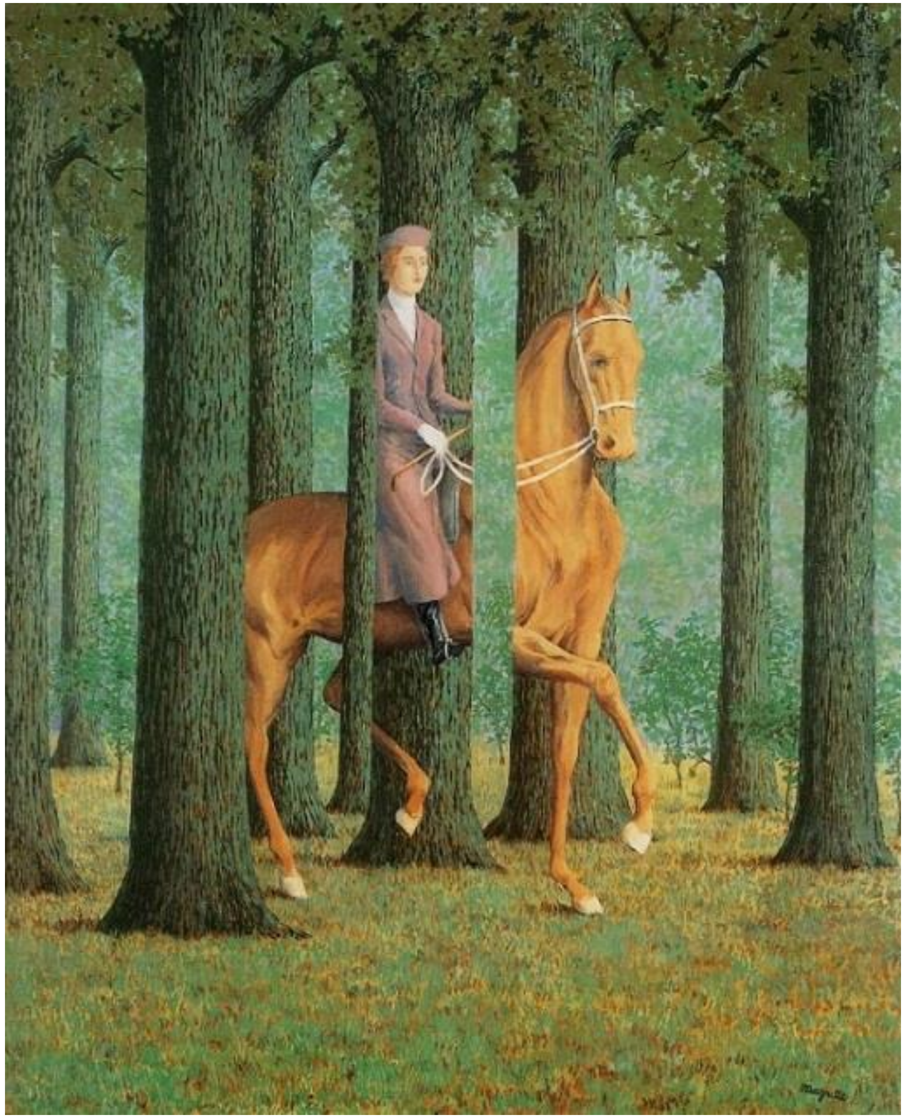
Motion



Figures from L. Zhang

http://www.brainconnection.com/teasers/?main=illusion/motion-shape

Occlusion

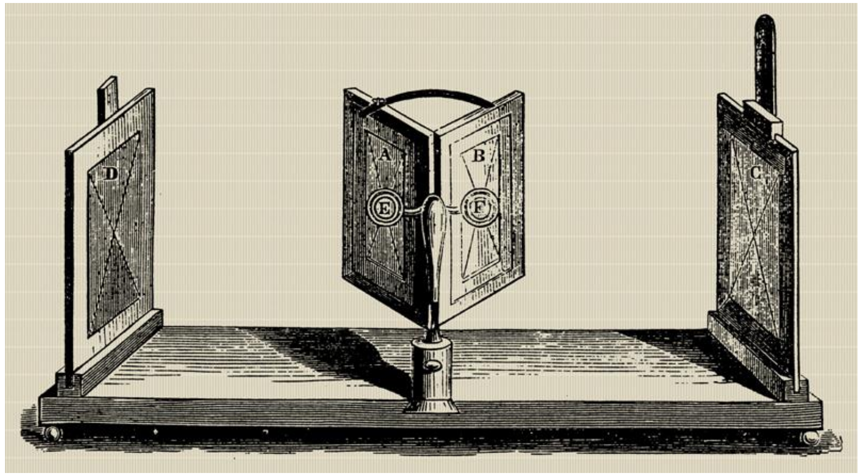


Rene Magritt'e famous painting *Le Blanc-Seing* (literal translation: "The Blank Signature") roughly translates as "free hand“. 1965

If stereo were critical for depth perception, navigation, recognition, etc., then this would be a problem

Stereo photography and stereo viewers

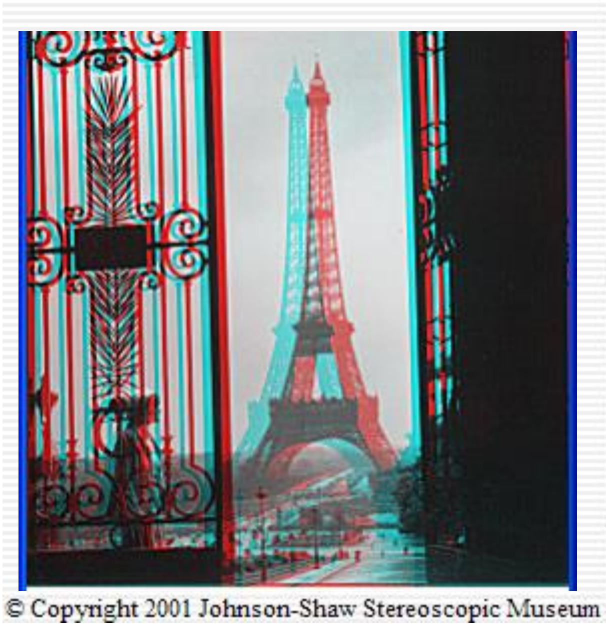
Take two pictures of the same subject from two slightly different viewpoints and display so that each eye sees only one of the images.



Invented by Sir Charles Wheatstone, 1838

Image from fisher-price.com

http://www.johnsonshawmuseum.org



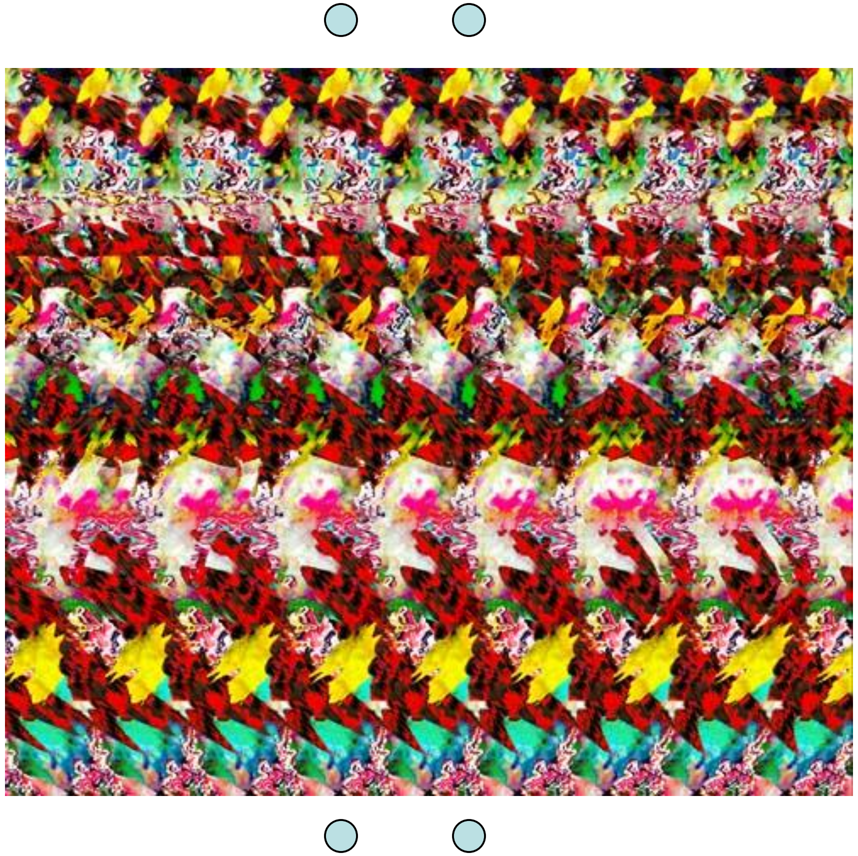
http://www.johnsonshawmuseum.org

Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923

http://www.well.com/~jimg/stereo/stereo_list.html

http://www.well.com/~jimg/stereo/stereo_list.html

Autostereograms



Exploit disparity as depth cue using single image.

(Single image random dot stereogram, Single image stereogram)

Images from magiceye.com

Autostereograms

Images from magiceye.com

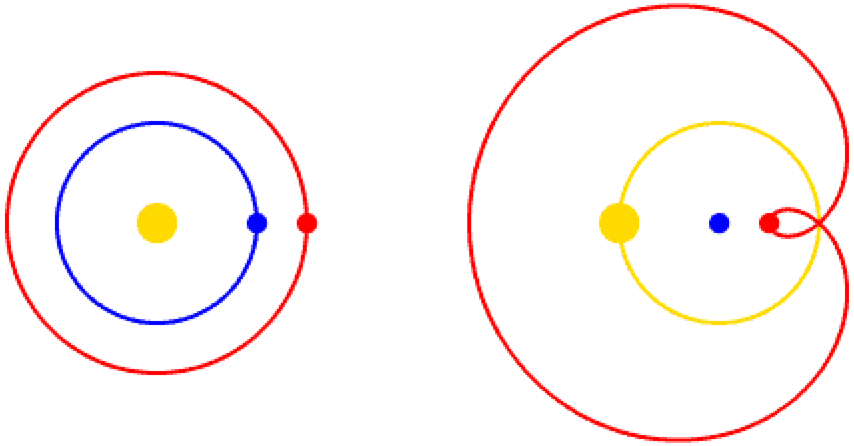
Parallax and our universe

Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilization, every king and peasant, every young couple in love, every mother and father, hopeful child, inventor and explorer, every teacher of morals, every corrupt politician, every "superstar," every "supreme leader," every saint and sinner in the history of our species lived there--on a mote of dust suspended in a sunbeam.

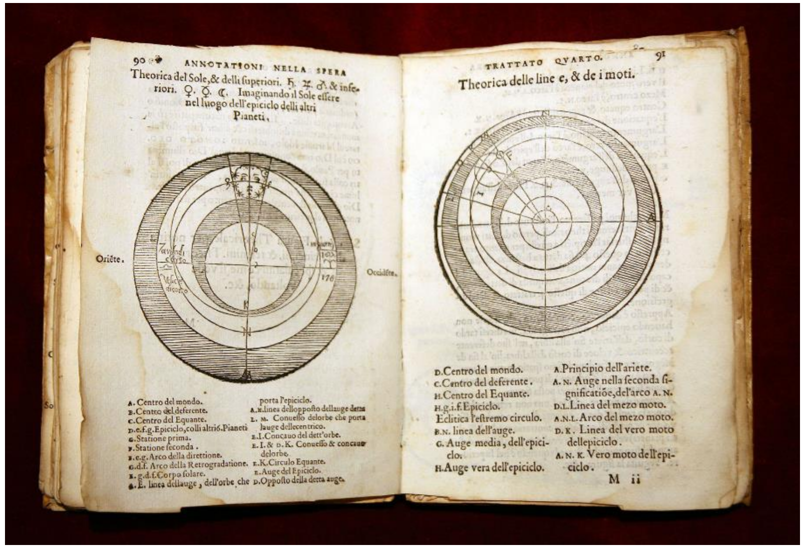
— Carl Sagan

https://en.wikipedia.org/wiki/Pale_Blue_Dot

Nicolaus Copernicus



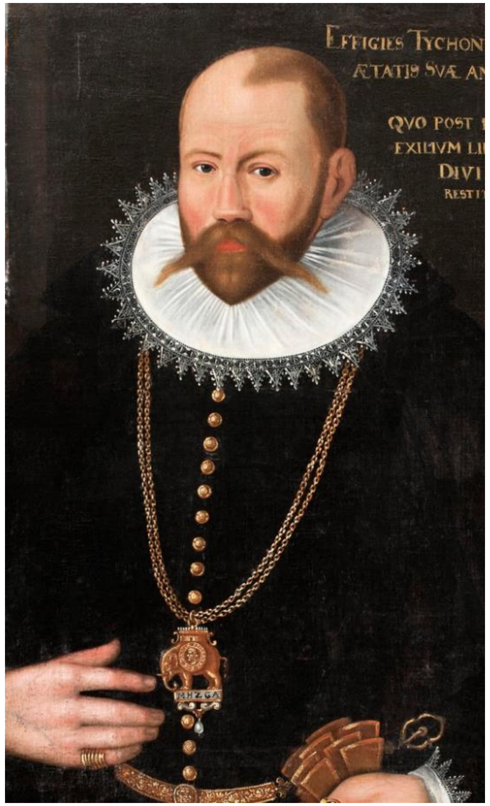
Motion of Sun (yellow), Earth (blue), and Mars (red). At left, Copernicus' heliocentric motion. At right, traditional geocentric motion, including the retrograde motion of Mars.



geocentric model (often exemplified specifically by the **Ptolemaic system**)

https://en.wikipedia.org/wiki/Heliocentrism

If the apparent motion of the planets is caused by parallax, why aren't we seeing parallax for stars?

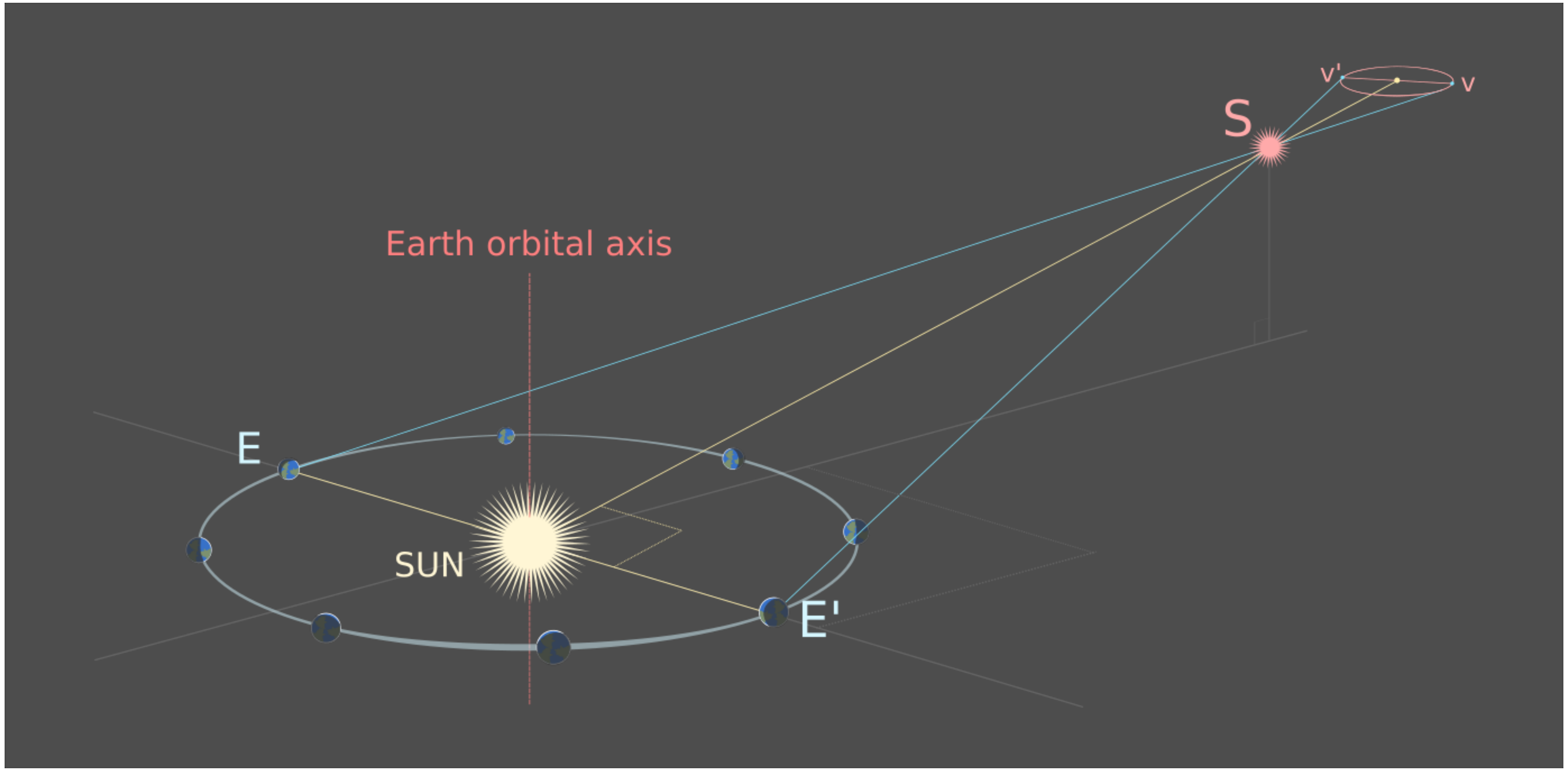


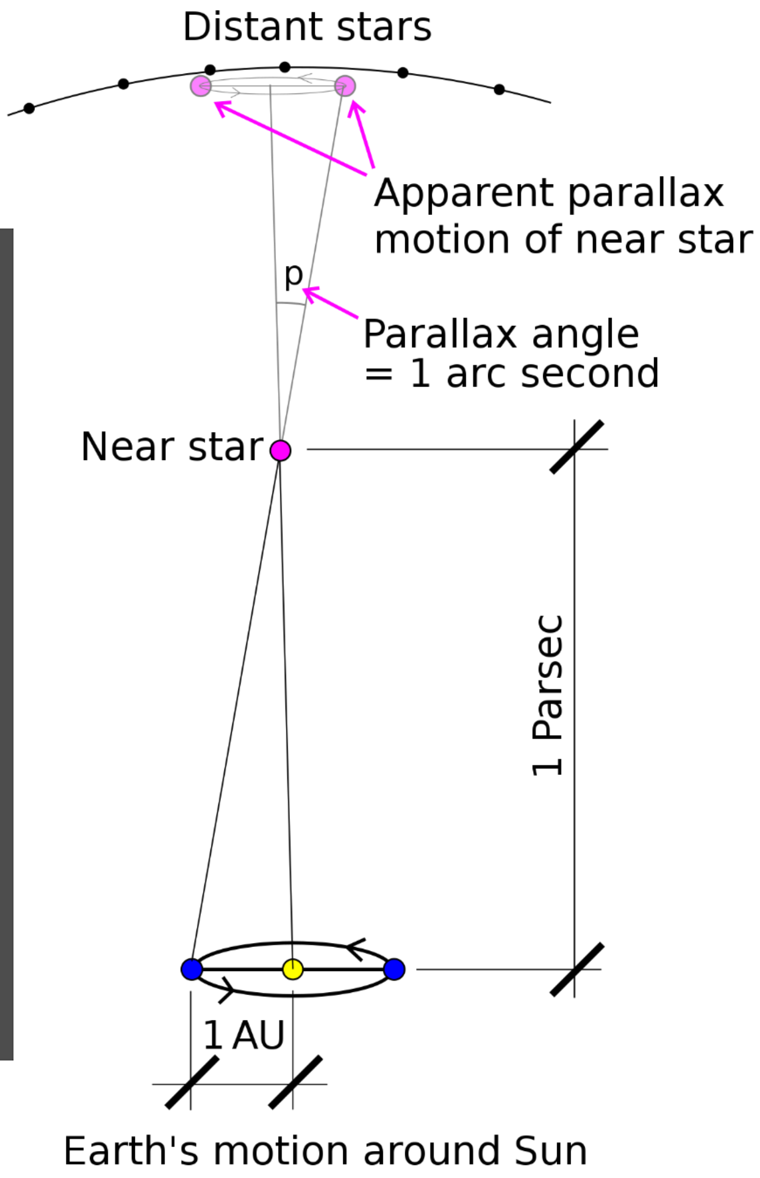
Tycho Brahe

it was one of Tycho Brahe's principal objections to Copernican heliocentrism that for it to be compatible with the lack of observable stellar parallax, there would have to be an enormous and unlikely void between the orbit of Saturn and the eighth sphere (the fixed stars).

The angles involved in these calculations are very small and thus difficult to measure. The nearest star to the Sun (and also the star with the largest parallax), Proxima Centauri, has a parallax of 0.7685 ± 0.0002 arcsec.[1] This angle is approximately that subtended by an object 2 centimeters in diameter located 5.3 kilometers away. First reliable measurements of parallax were not made until 1838, by Friedrich Bessel

https://en.wikipedia.org/wiki/Stellar_parallax





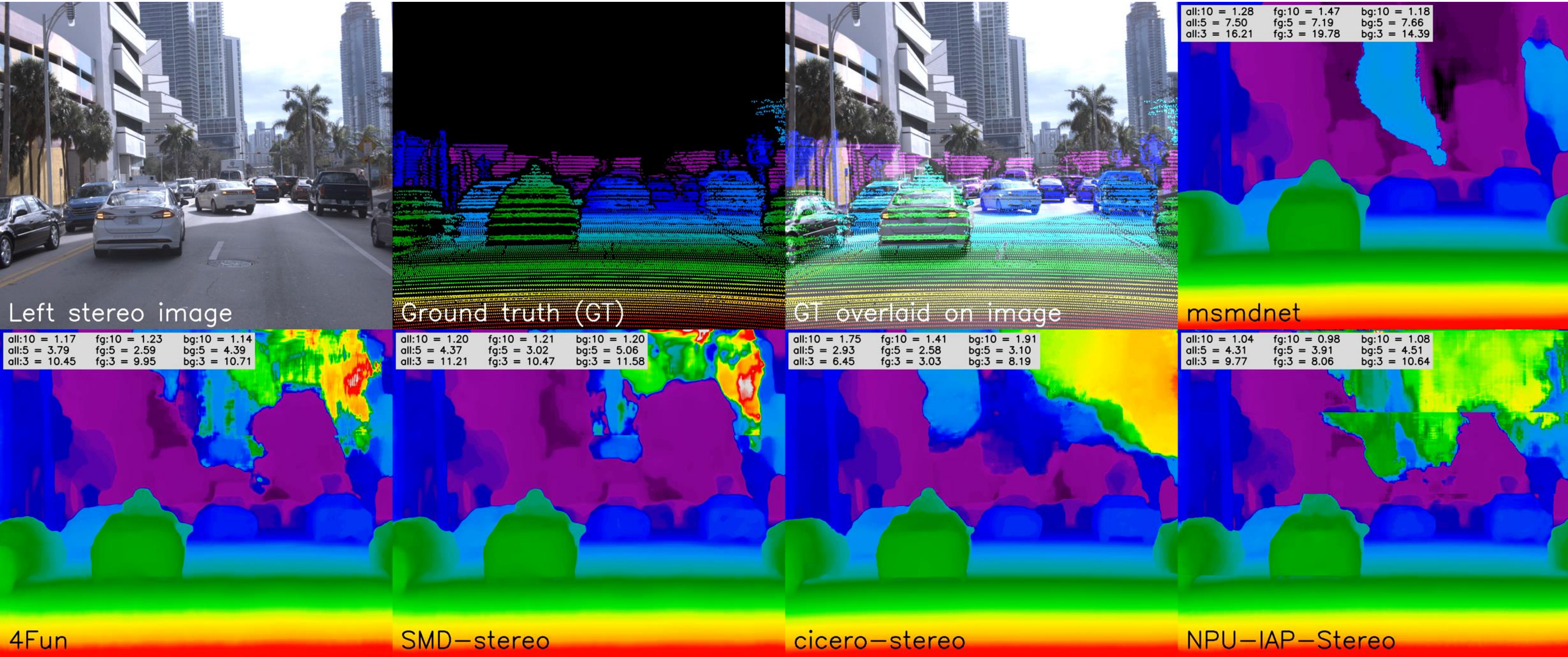
https://en.wikipedia.org/wiki/Stellar_parallax

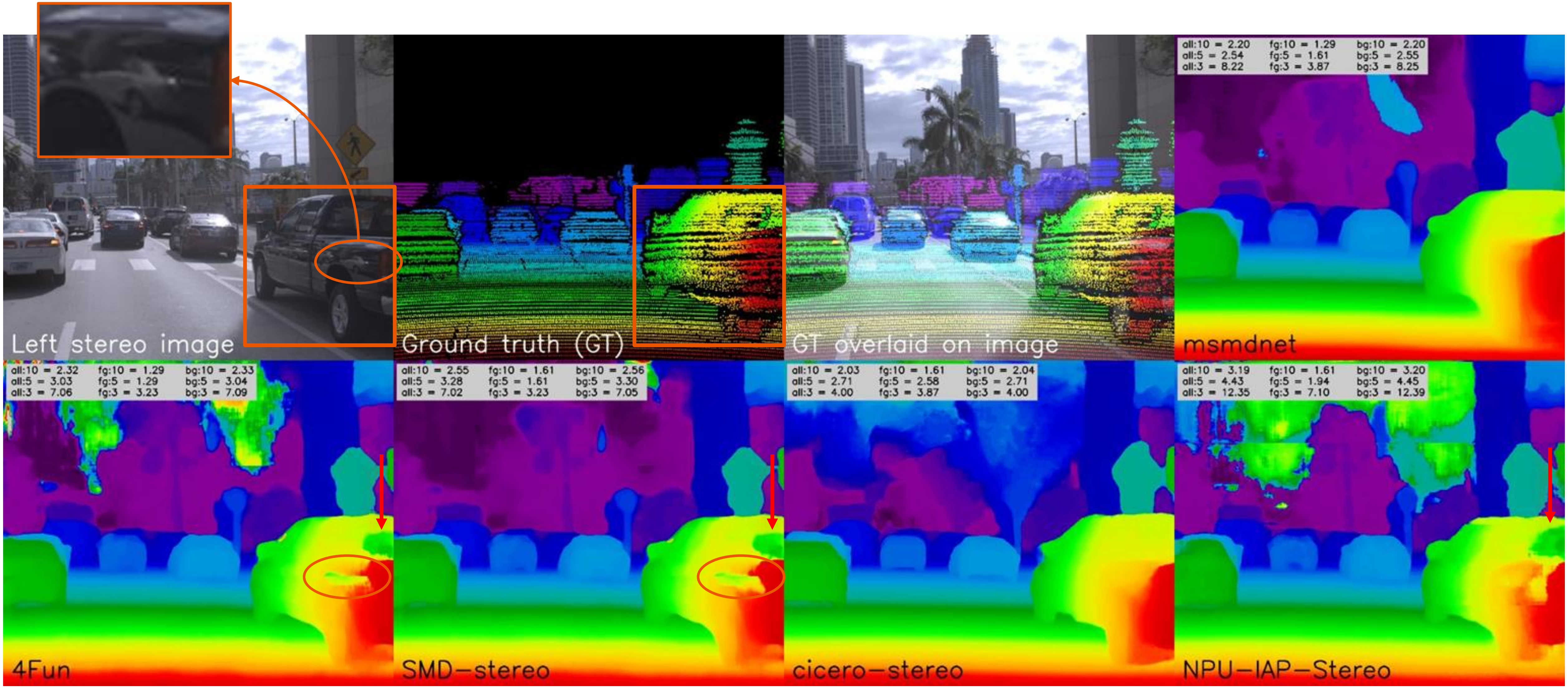
Stereo vision

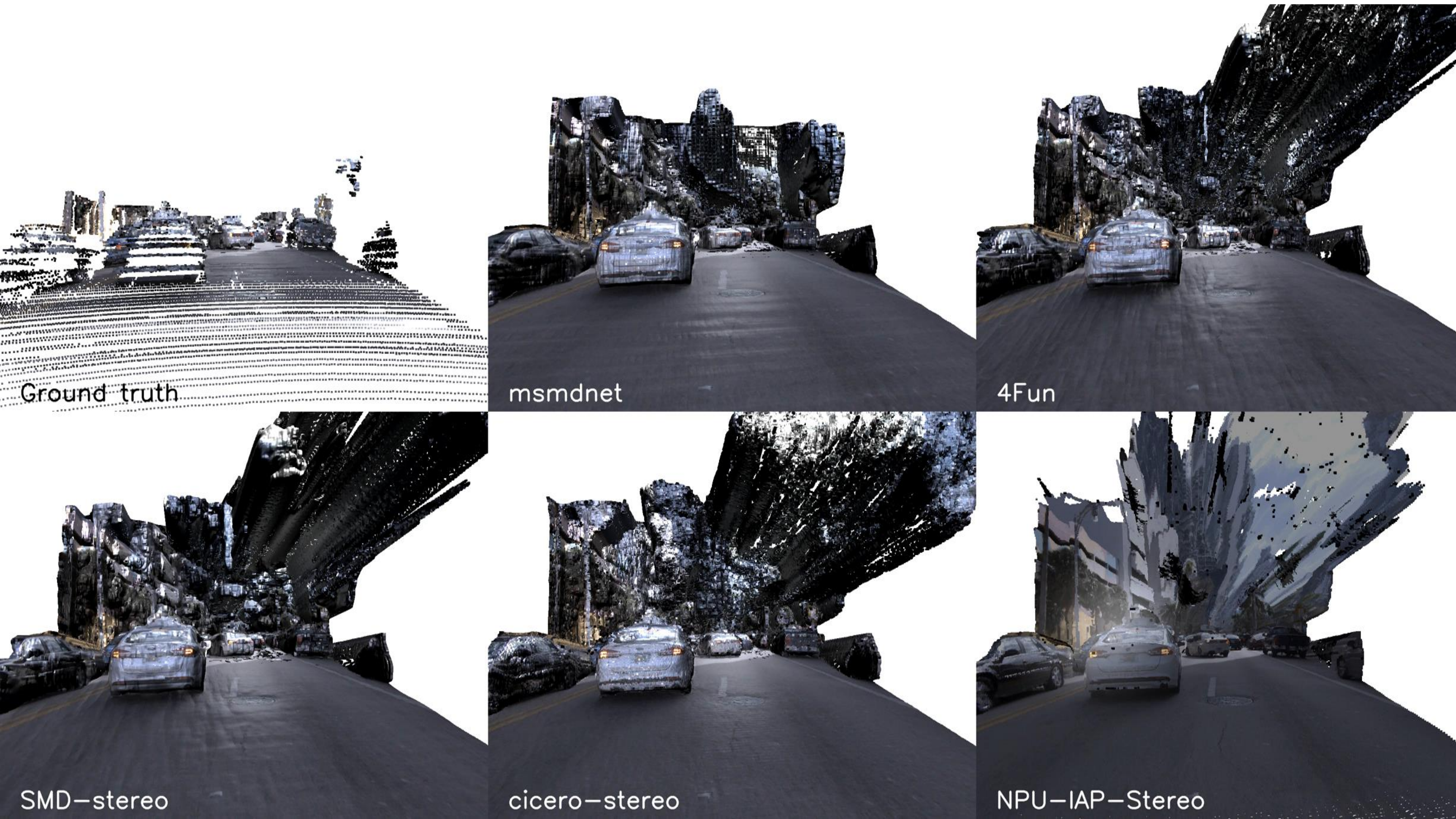
Two cameras, simultaneous views

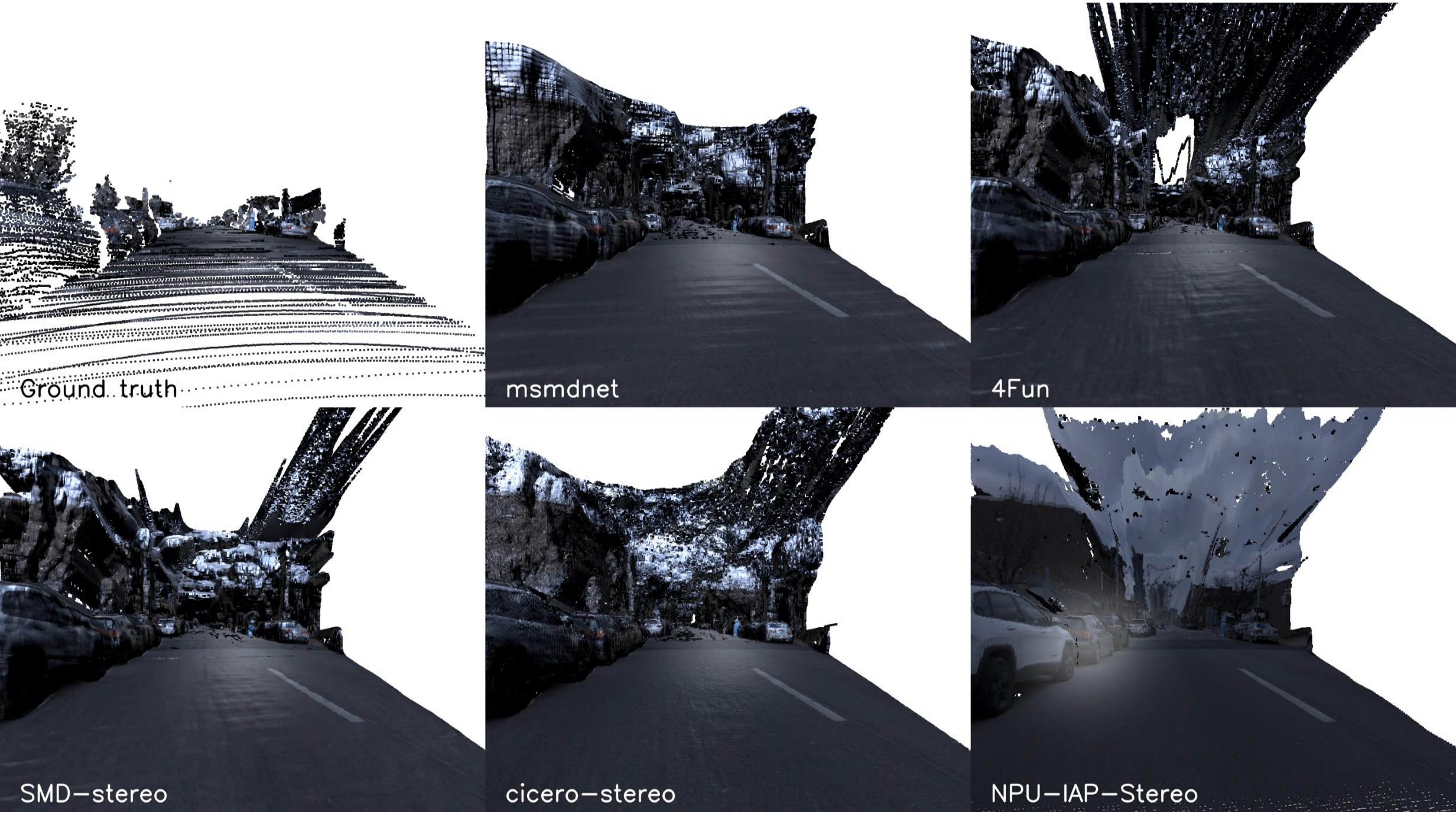
Single moving camera and static scene

Modern stereo depth estimation example



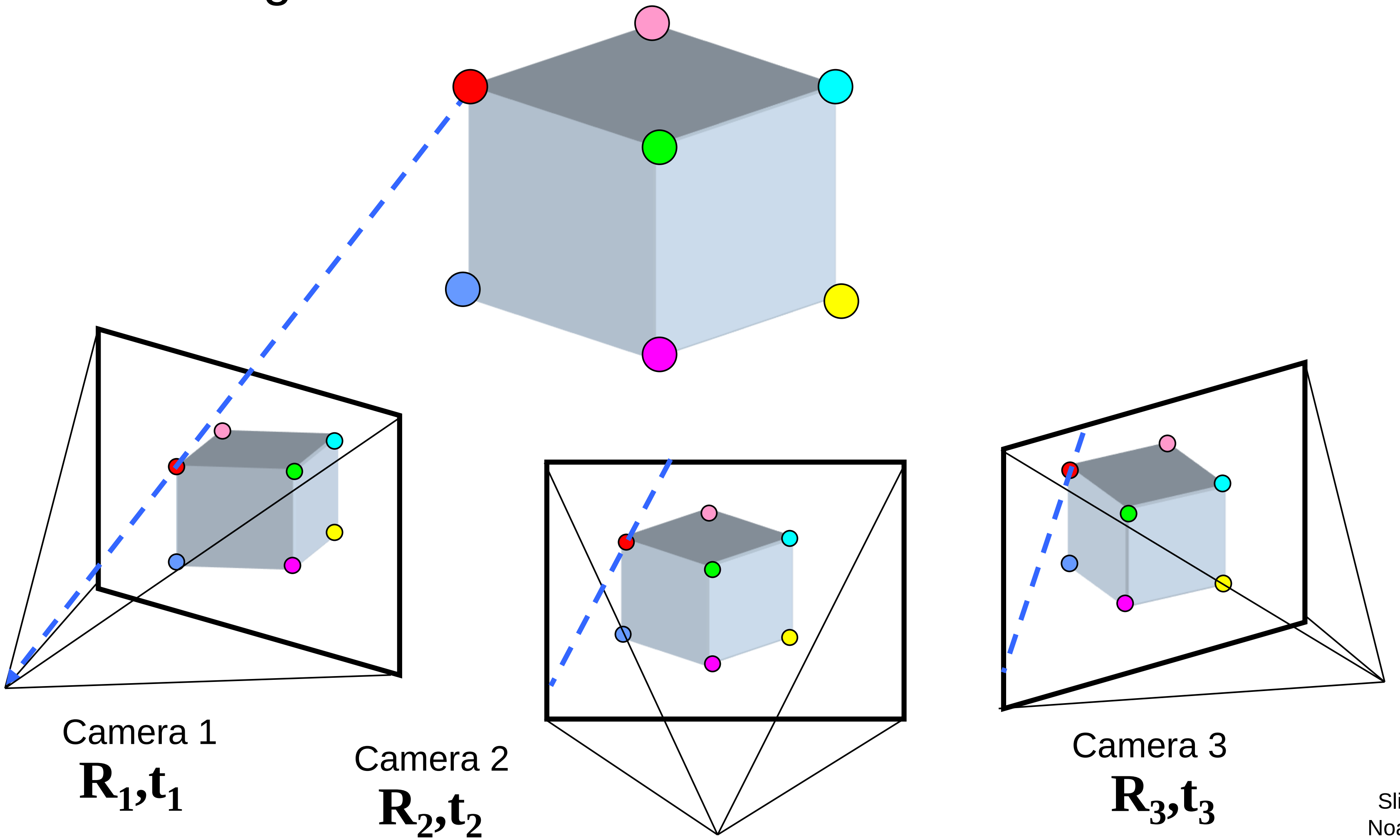

Left stereo image
Ground truth (GT)
GT overlaid on image
all:10 = 2.20 fg:10 = 1.29 bg:10 = 2.20
all:5 = 2.54 fg:5 = 1.61 bg:5 = 2.55
all:3 = 8.22 fg:3 = 3.87 bg:3 = 8.25
msmdnet
all:10 = 2.32 fg:10 = 1.29 bg:10 = 2.33
all:5 = 3.03 fg:5 = 1.29 bg:5 = 3.04
all:3 = 7.06 fg:3 = 3.23 bg:3 = 7.09
4Fun
all:10 = 2.55 fg:10 = 1.61 bg:10 = 2.56
all:5 = 3.28 fg:5 = 1.61 bg:5 = 3.30
all:3 = 7.02 fg:3 = 3.23 bg:3 = 7.05
SMD-stereo
all:10 = 2.03 fg:10 = 1.61 bg:10 = 2.04
all:5 = 2.71 fg:5 = 2.58 bg:5 = 2.71
all:3 = 4.00 fg:3 = 3.87 bg:3 = 4.00
cicero-stereo
all:10 = 3.19 fg:10 = 1.61 bg:10 = 3.20
all:5 = 4.43 fg:5 = 1.94 bg:5 = 4.45
all:3 = 12.35 fg:3 = 7.10 bg:3 = 12.39
NPU-IAP-Stereo


Ground truth
msmdnet
4Fun
SMD-stereo
cicero-stereo
NPU-IAP-Stereo


Ground truth
msmdnet
4Fun
SMD-stereo
cicero-stereo
NPU-IAP-Stereo

Multi-view geometry problems

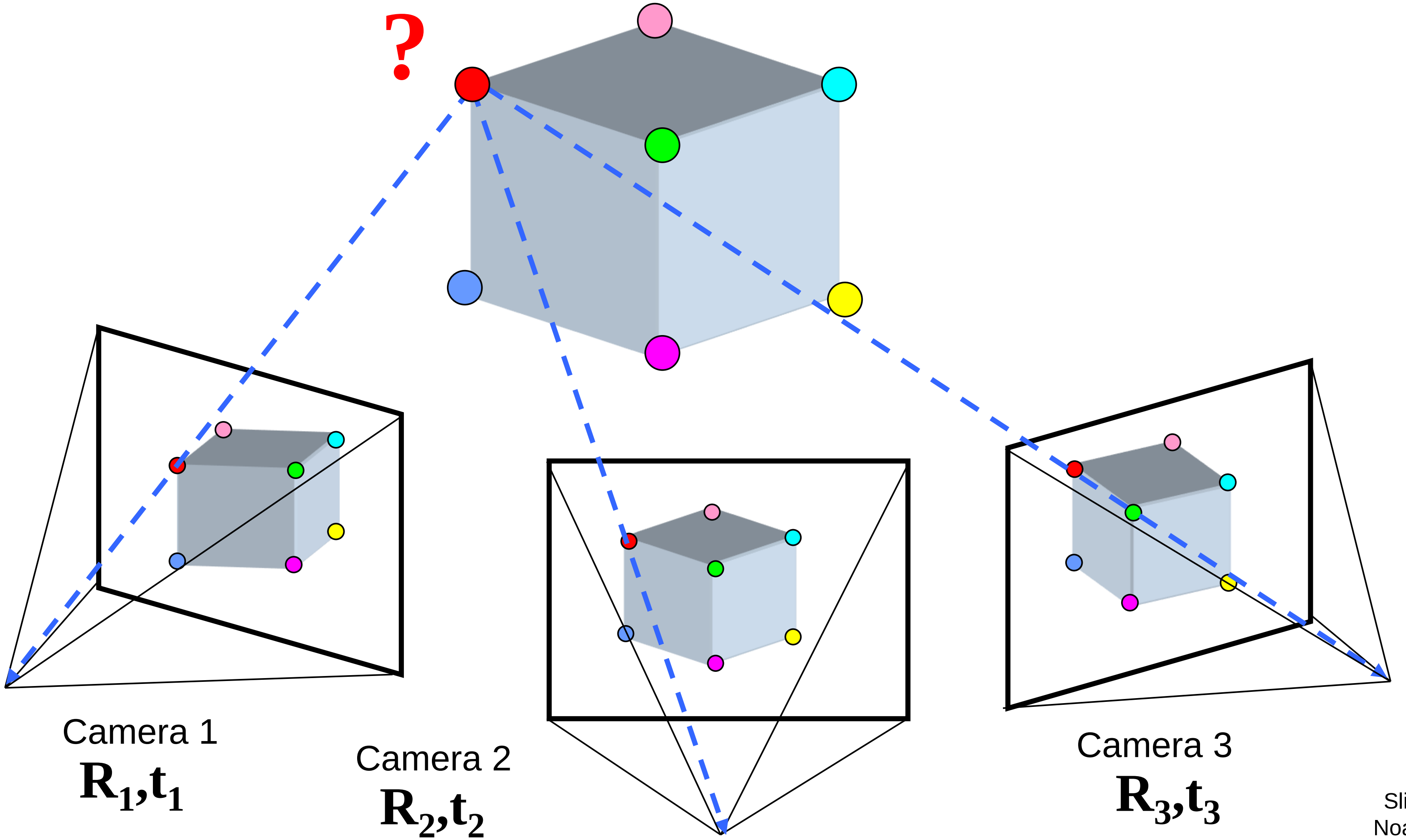
- **Stereo correspondence:** Given a point in one of the images, where could its corresponding points be in the other images?



Slide credit: Noah Snavely

Multi-view geometry problems

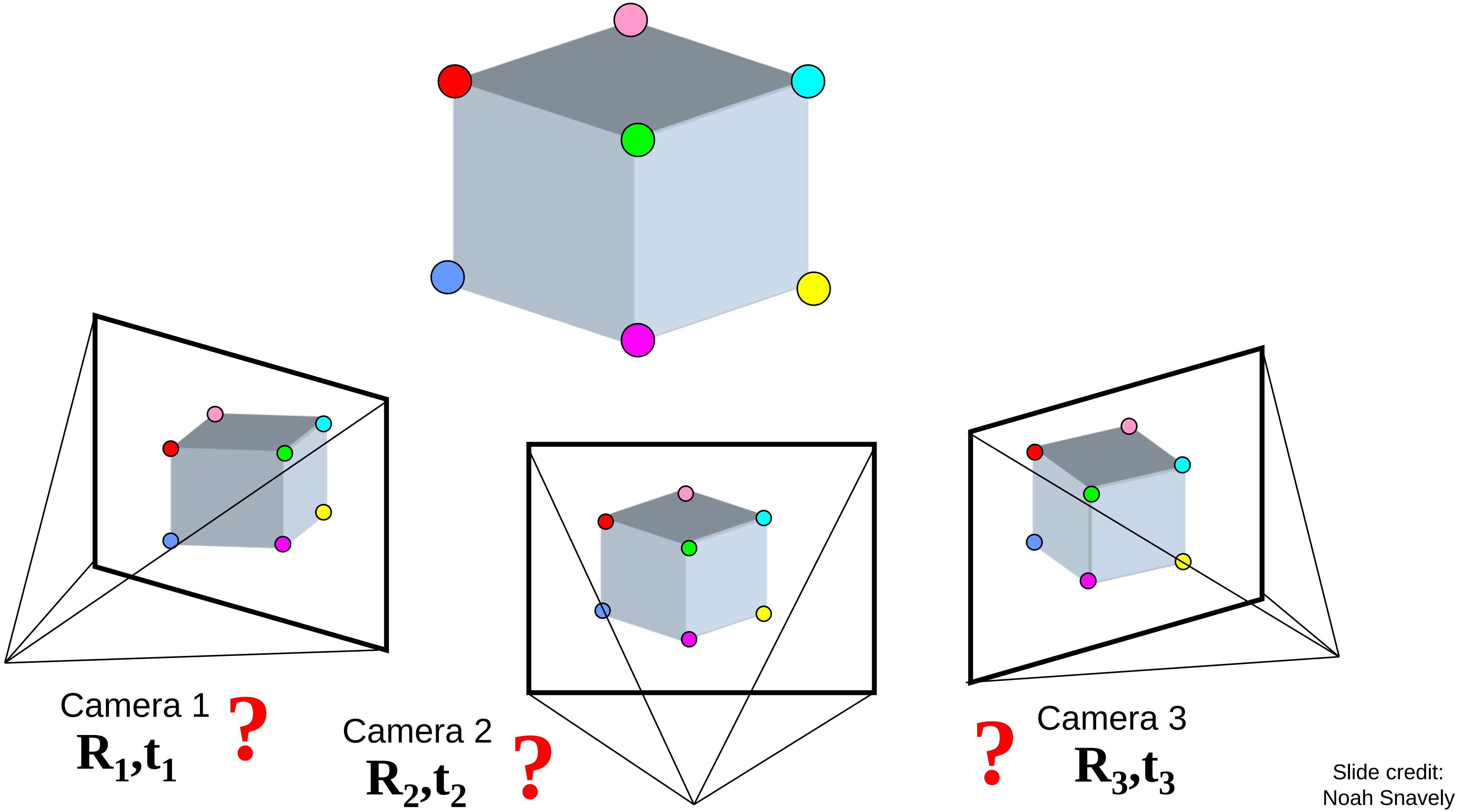
- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point



Slide credit: Noah Snavely

Multi-view geometry problems

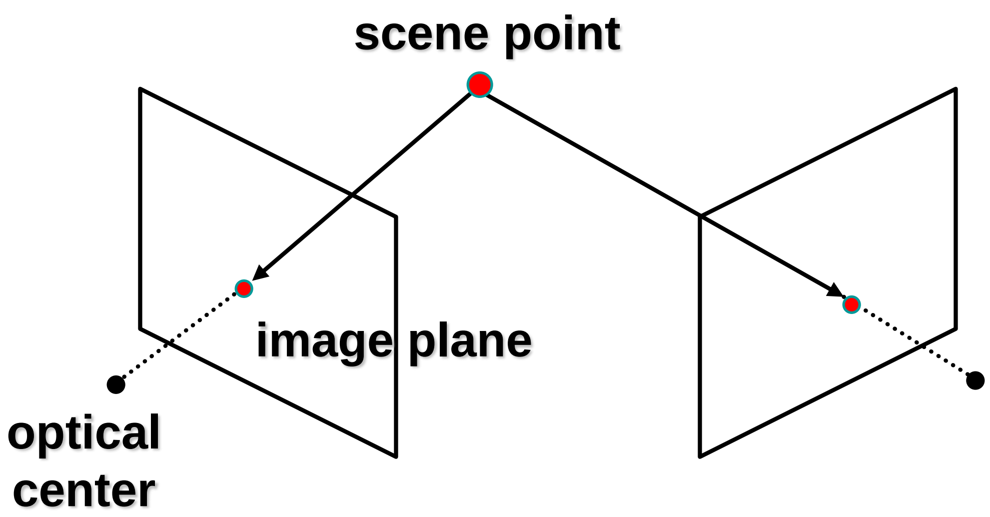
- **Motion:** Given a set of corresponding points in two or more images, compute the camera parameters

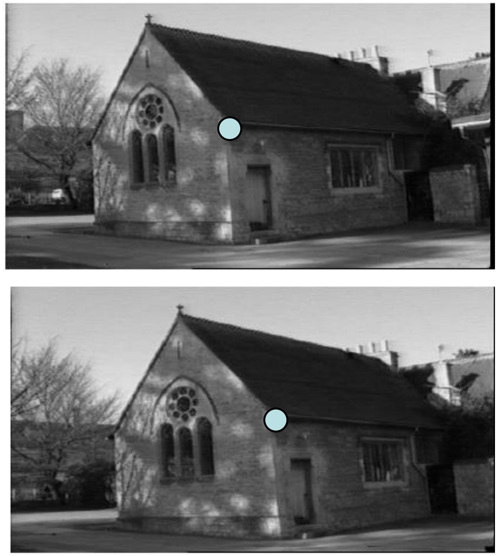


Slide credit: Noah Snavely

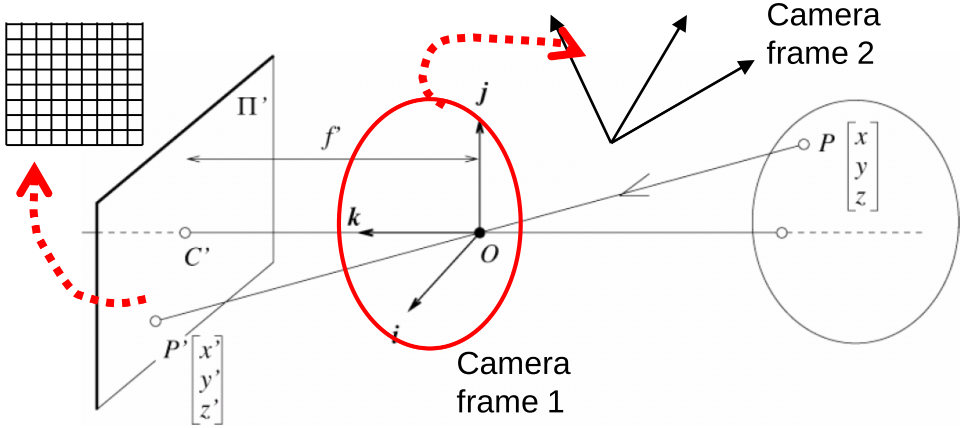
Estimating depth with stereo

- **Stereo**: shape from "motion" between two views
- We'll need to consider:
  - Info on camera pose ("calibration")
  - Image point correspondences





Camera parameters



Extrinsic parameters:
Camera frame 1 ←→ Camera frame 2

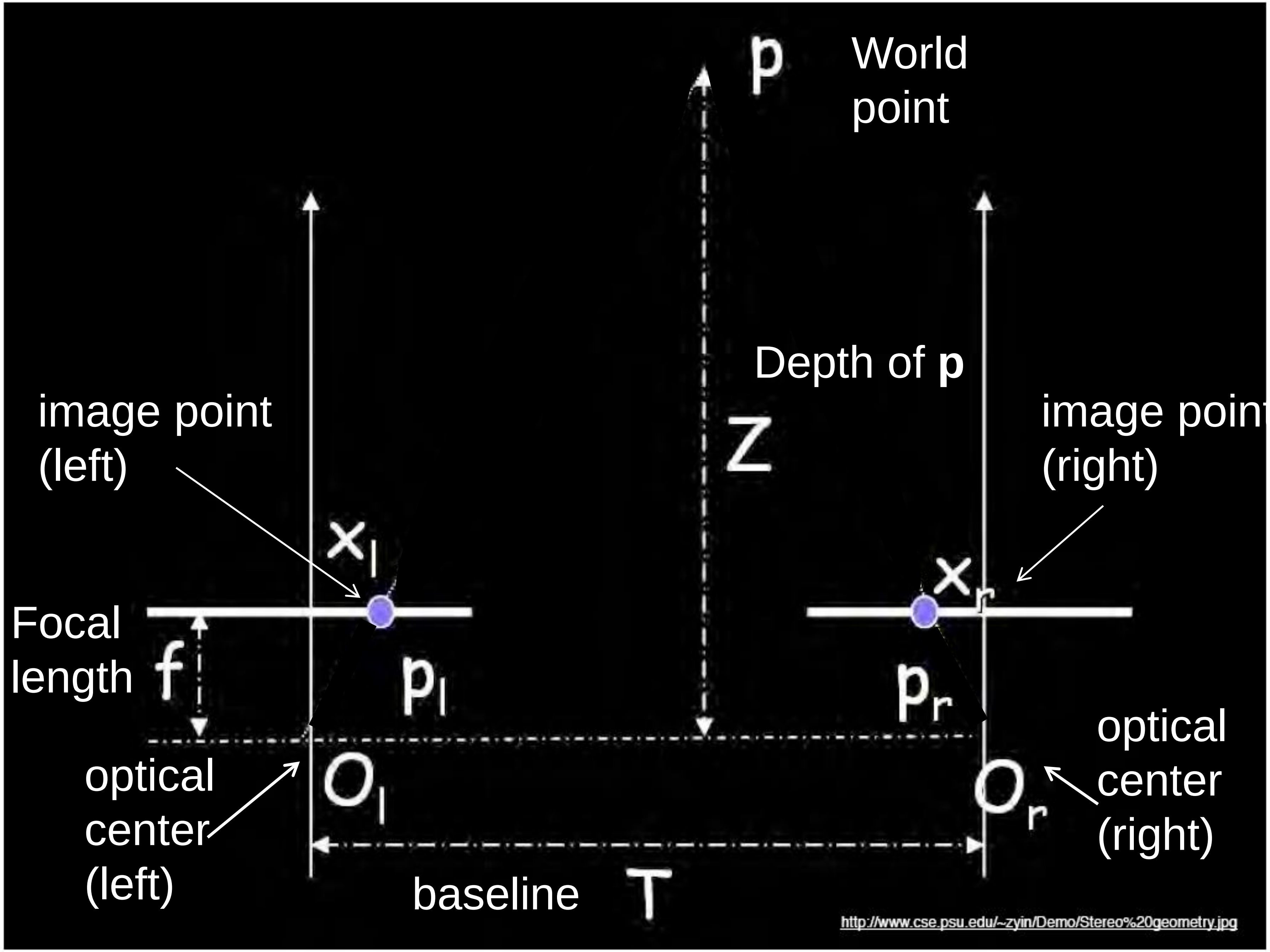
Intrinsic parameters:
Image coordinates relative to camera ←→ Pixel coordinates

- *Extrinsic* params: rotation matrix and translation vector
- *Intrinsic* params: focal length, pixel sizes (mm), image center point, radial distortion parameters

We'll assume for now that these parameters are given and fixed.

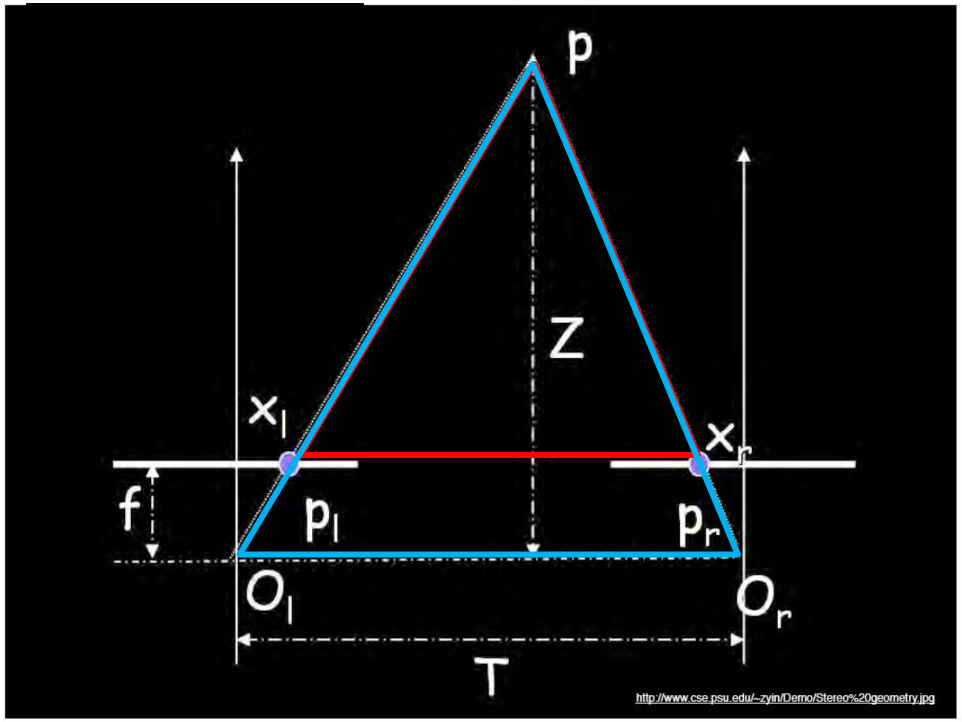
Geometry for a simple stereo system

- First, assuming parallel optical axes, known camera parameters (i.e., calibrated cameras):


p
World
point
Depth of **p**
Z
image point
(left)
image point
(right)
x_l
x_r
Focal
length
f
p_l
p_r
optical
center
(left)
O_l
O_r
optical
center
(right)
baseline
T
http://www.cse.psu.edu/~zyin/Demo/Stereo%20geometry.jpg

Geometry for a simple stereo system

- Assume parallel optical axes, known camera parameters (i.e., calibrated cameras). **What is expression for Z?**



Similar triangles (p_l, P, p_r) and (O_l, P, O_r):

$$\frac{T - x_l + x_r}{Z - f} = \frac{T}{Z}$$

$$Z = f\,\frac{T}{x_l - x_r}$$

disparity → $x_l - x_r$

Depth from disparity

image I(x,y)

Disparity map D(x,y)

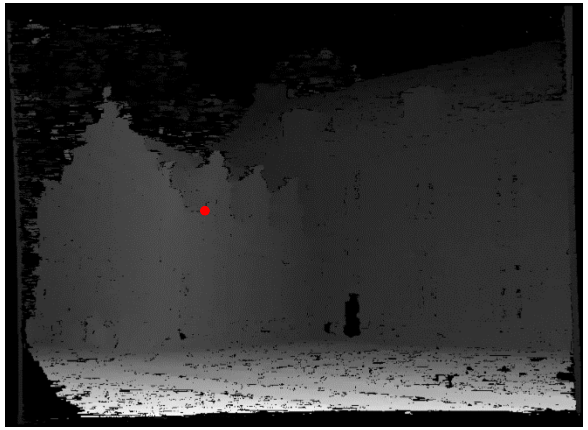
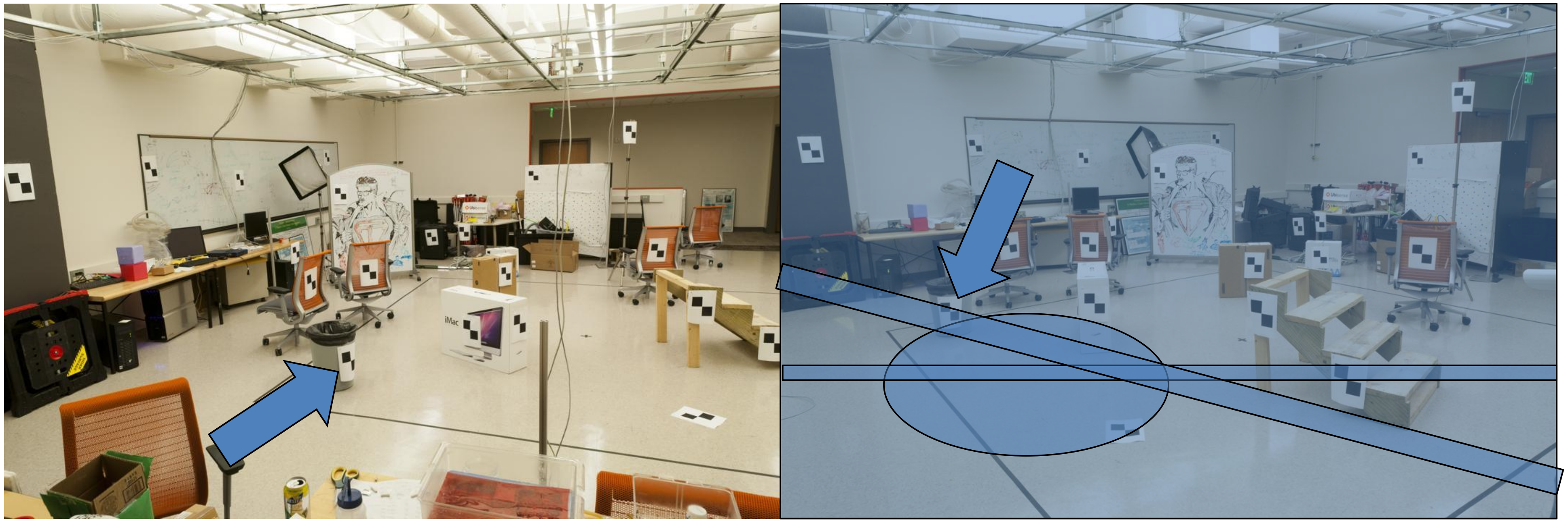


image I´(x´,y´)

$$(x',y')=(x+D(x,y),\ y)$$

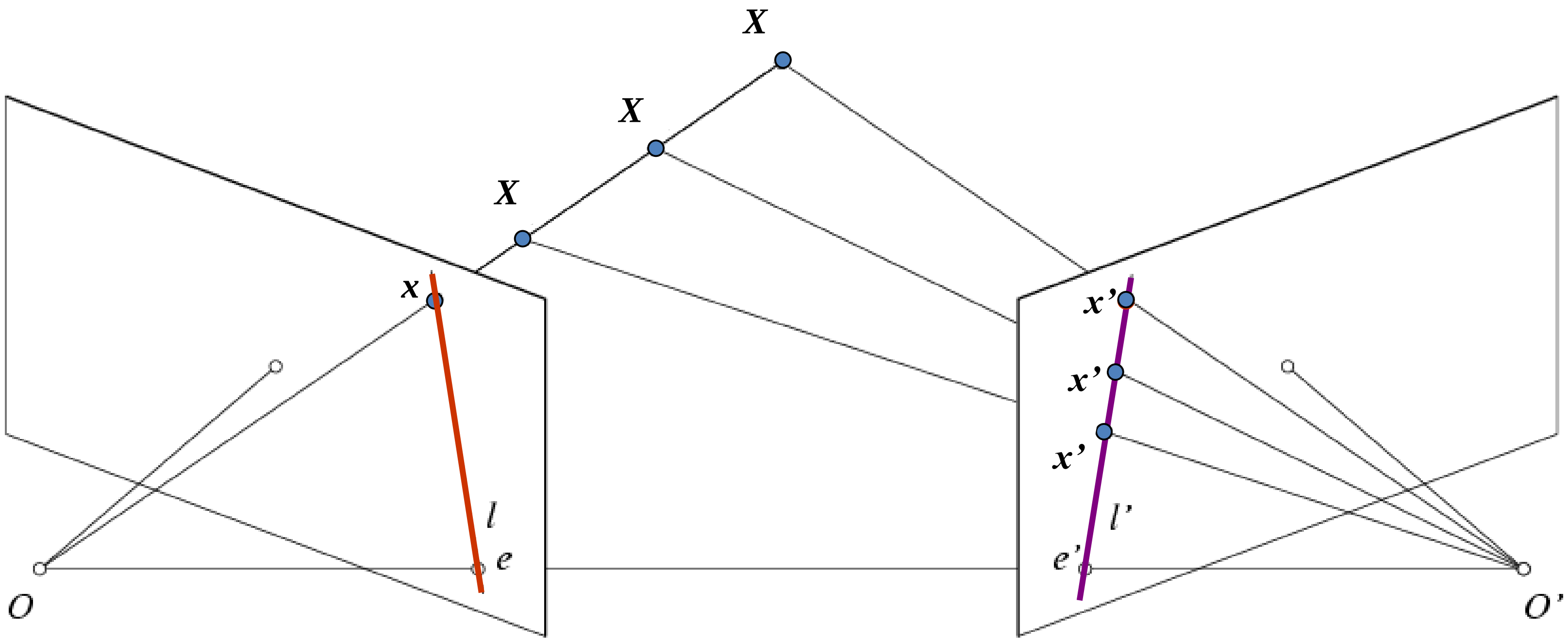
So if we could find the **corresponding points** in two images, we could **estimate relative depth**…

Where do we need to search?



Key idea: Epipolar constraint

Key idea: Epipolar constraint

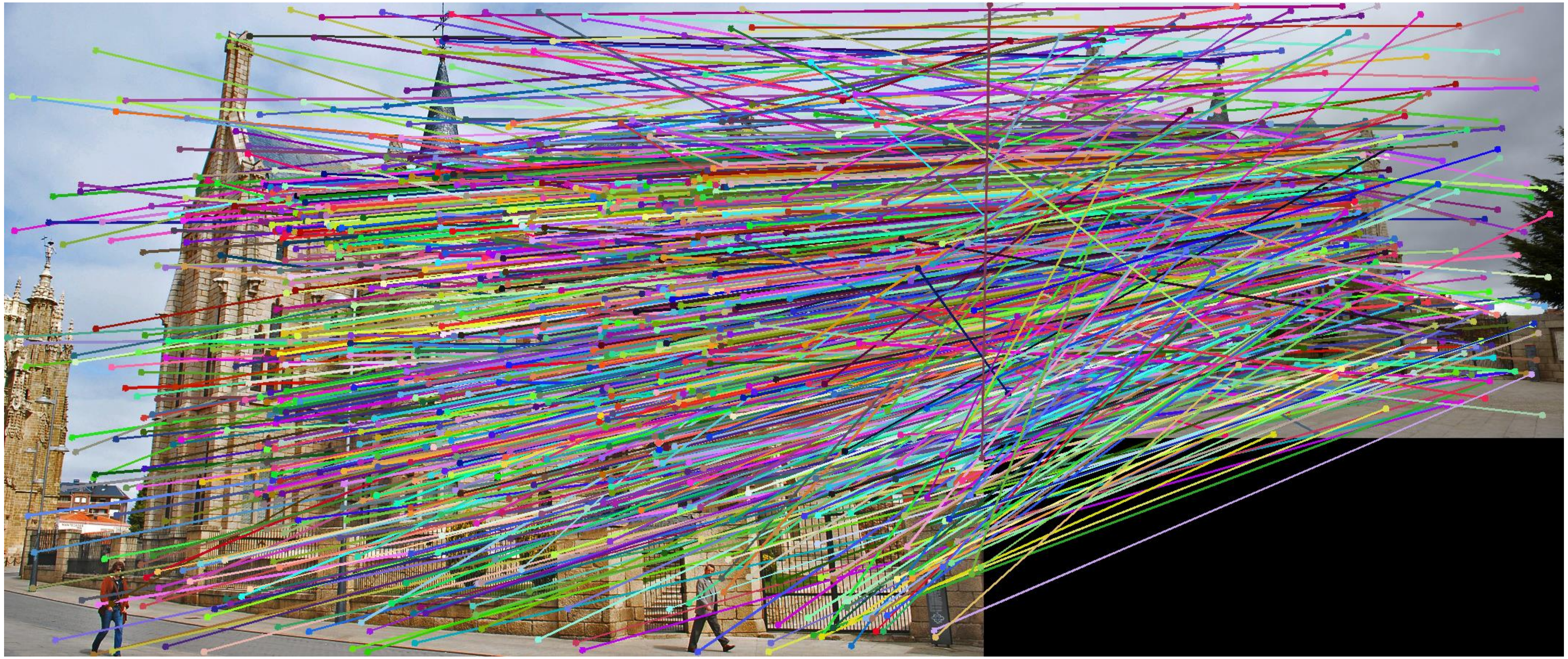


Potential matches for *x* have to lie on the corresponding line *l'*.

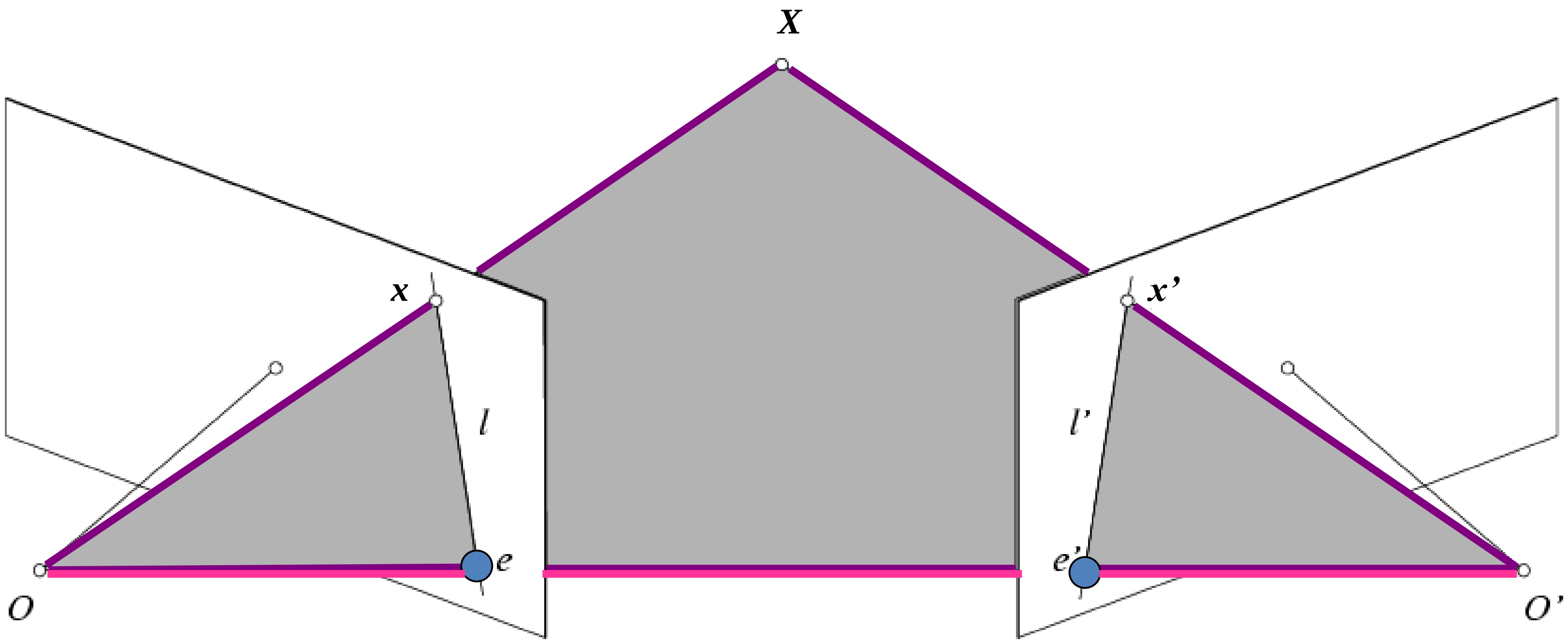
Potential matches for *x'* have to lie on the corresponding line *l*.

Wouldn't it be nice to know where matches can live? To constrain our 2d search to 1d.

VLFeat's 800 most confident matches among 10,000+ local features.

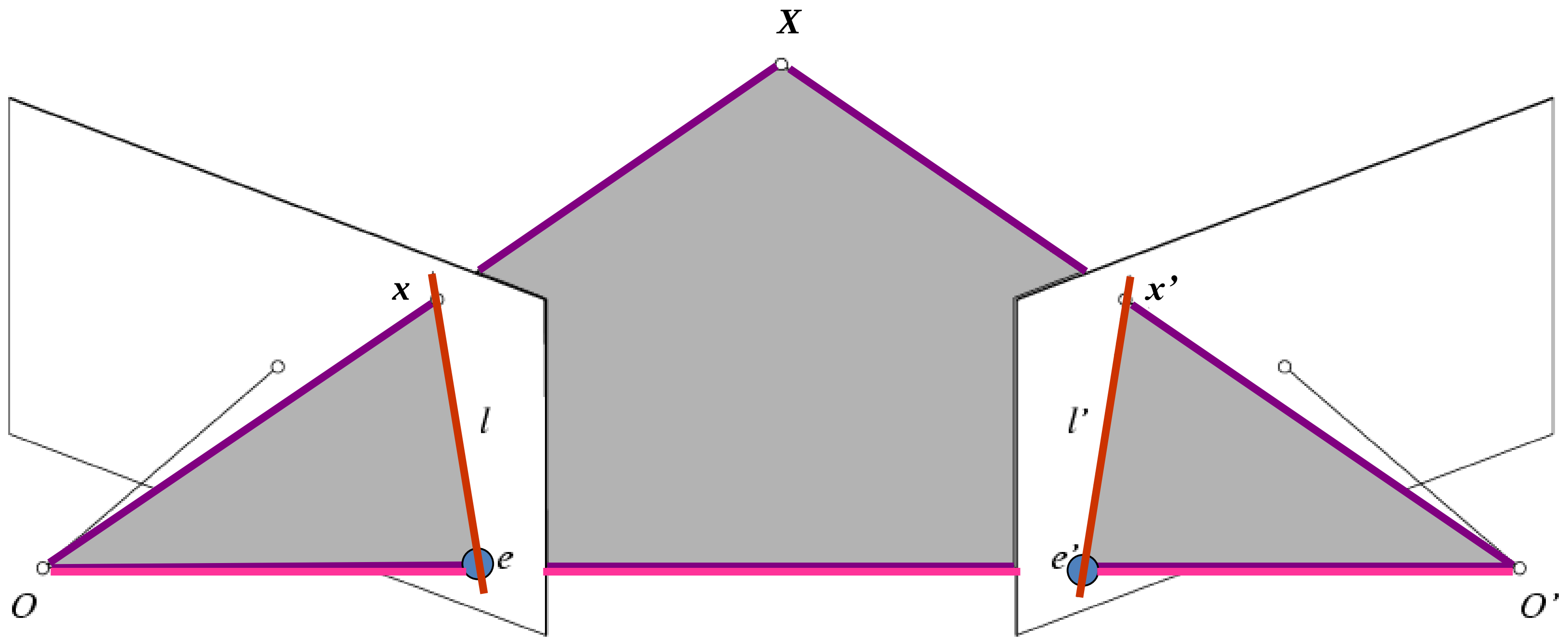


Epipolar geometry: notation



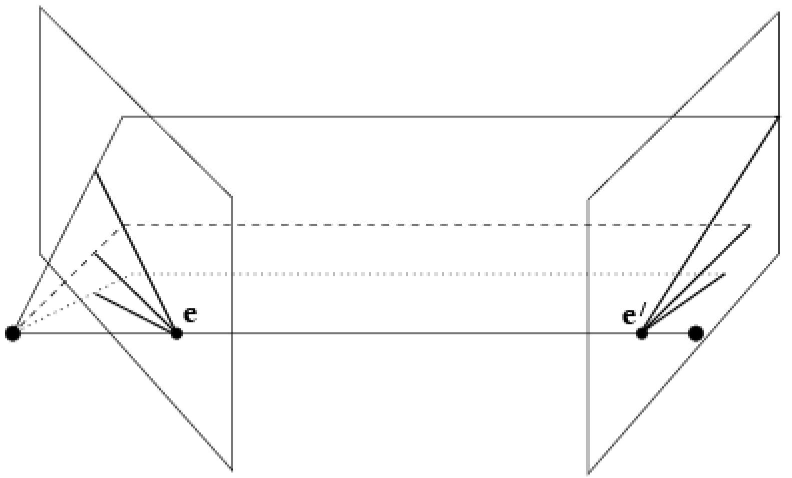
- **Baseline** – line connecting the two camera centers
- **Epipoles**
  = intersections of baseline with image planes
  = projections of the other camera center
- **Epipolar Plane** – plane containing baseline (1D family)

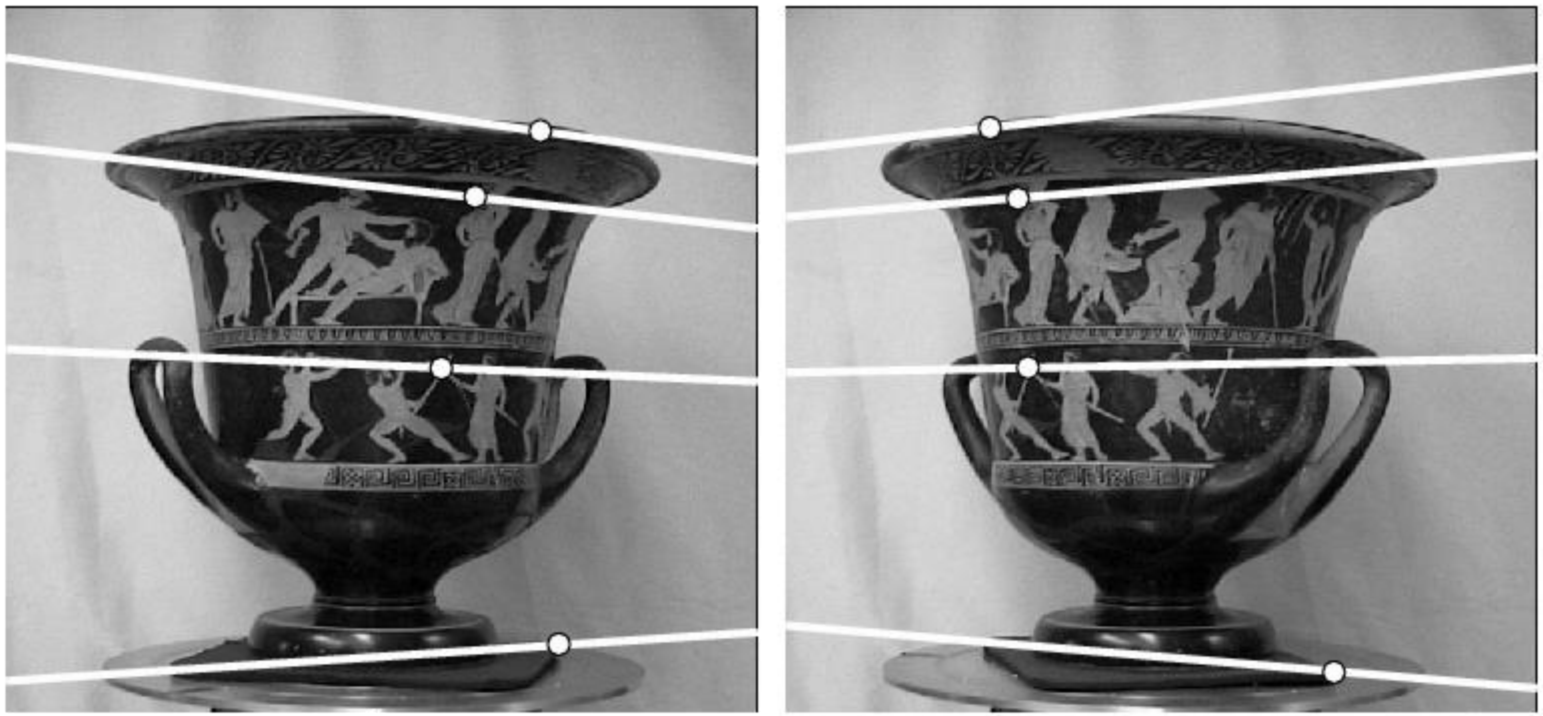
Epipolar geometry: notation



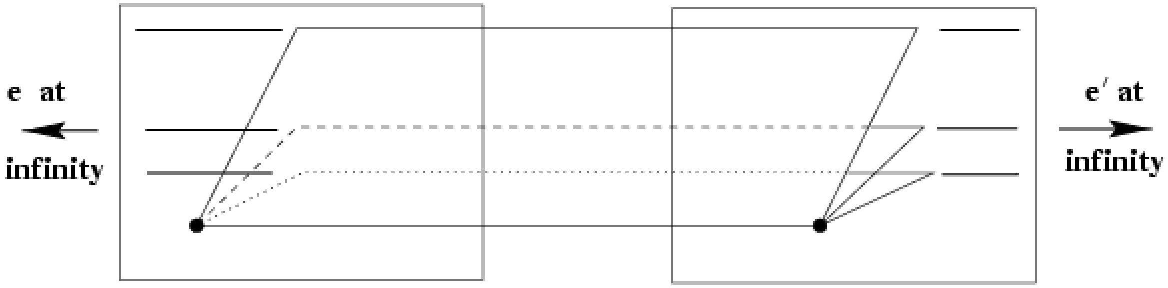
- **Baseline** – line connecting the two camera centers
- **Epipoles**
  = intersections of baseline with image planes
  = projections of the other camera center
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)

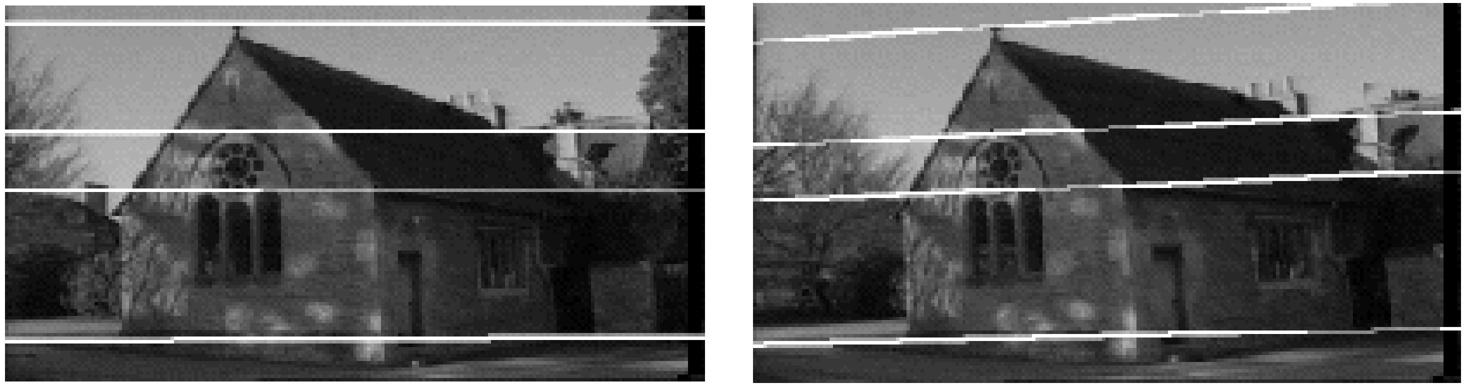
Example: Converging cameras





Example: Motion parallel to image plane

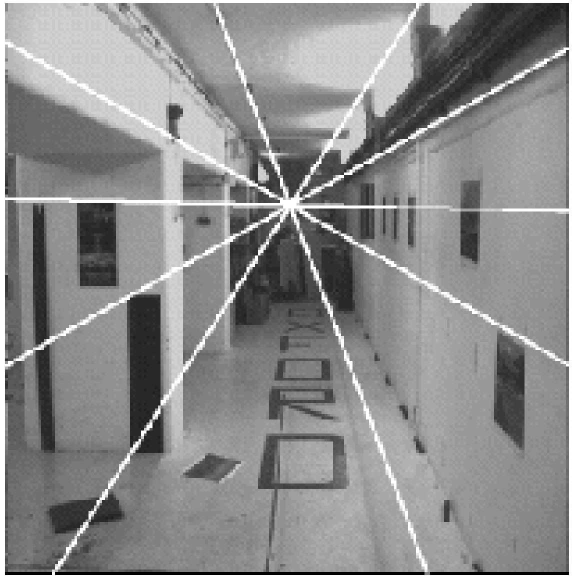


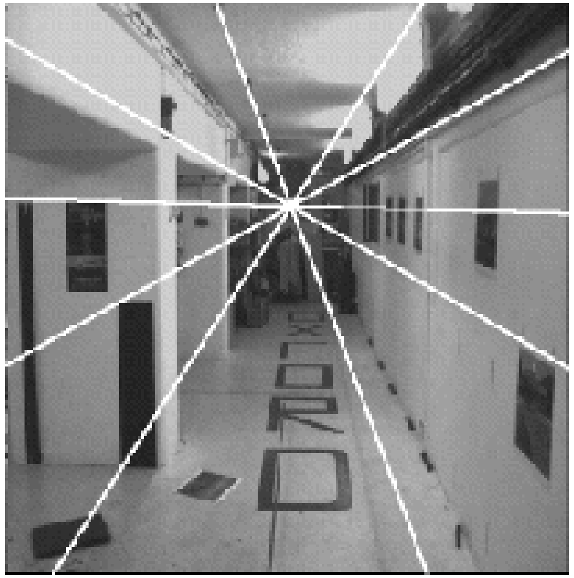


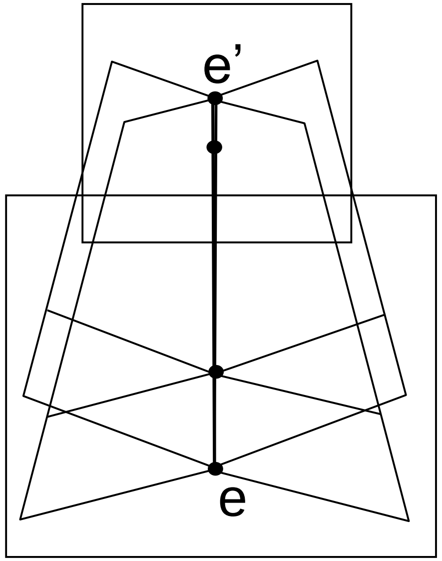
Example: Forward motion

What would the epipolar lines look like if the camera moves directly forward?

Example: Forward motion

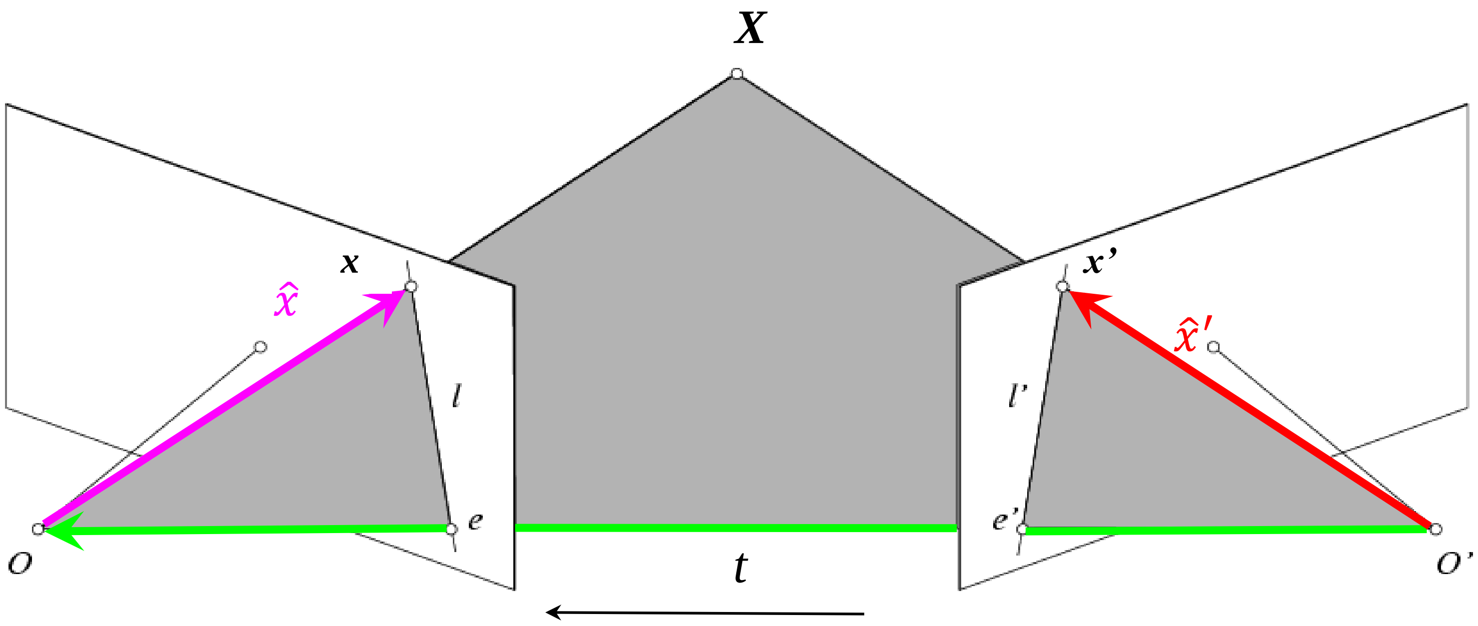






Epipole has same coordinates in both images.
Points move along lines radiating from e: "Focus of expansion"

Epipolar constraint: Calibrated case



$$\hat{x} = K^{-1}x = X \qquad \hat{x}' = K'^{-1}x' = X'$$

$$\hat{x} \cdot [t \times (R\hat{x}')] = 0$$

(because $\hat{x}, R\hat{x}'$, and t are co-planar)

To be continued